

09/913494

PC/EP 00/00978

BUNDESREPUBLIK DEUTSCHLAND

EP 00 / 078

PRIORITY
DOCUMENTSUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

REC'D 01 MAR 2000	
WIPO	PCT

Bescheinigung

Die Merck Patent GmbH in Darmstadt/Deutschland hat eine Patentanmeldung
unter der Bezeichnung

"Glucose-Dehydrogenase-Fusionsproteine und ihre Verwendung
in Expressionssystemen"

am 19. Februar 1999 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprüng-
lichen Unterlagen dieser Patentanmeldung.

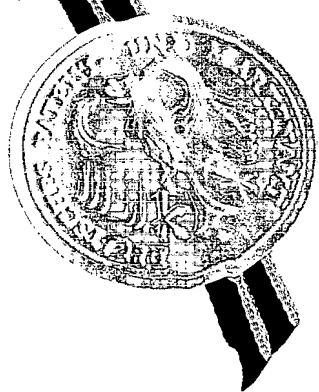
Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig die Symbole
C 12 N und C 12 Q der Internationalen Patentklassifikation erhalten.

München, den 8. November 1999
Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Seiler

Aktenzeichen: 199 06 920.4

**Merck Patent Gesellschaft
mit beschränkter Haftung
64271 Darmstadt**

**Glucose-Dehydrogenase-Fusionsproteine und ihre
Verwendung in Expressionssystemen**

Glucose-Dehydrogenase-Fusionsproteine und ihre Verwendung in Expressionssystemen

Die Erfindung betrifft neue rekombinante Fusionsproteine, welche als ein Be-
5 standteil eine Proteinsequenz mit der biologischen Aktivität von Glucose-
Dehydrogenase (GlcDH) enthalten sowie ihre Verwendung zum einfachen und
effizienten Nachweis von beliebigen, vorzugsweise als Fusionspartner dienenden
Proteinen / Polypeptiden bzw. zur raschen Optimierung von Expressionssystemen,
welche besagte Proteine / Polypeptide zu exprimieren in der Lage sind.

10 Dabei übernimmt die GlcDH, bzw die Sequenz, die die biologische Aktivität von
GlcDH aufweist, die Rolle eines Marker- bzw. Detektorproteins. Dieses Enzym
besitzt als Besonderheit eine außerordentliche Stabilität gegenüber denaturie-
renden Agenzien wie SDS. GlcDH als Marker- bzw. Detektorprotein zeigt selbst
15 nach den reduzierenden und denaturierenden Bedingungen von SDS-PAGE-
Gelen eine unverminderte enzymatische Aktivität. Fusionsproteine, die GlcDH
enthalten, sind daher mit einer auf diesem überraschenden Verhalten beruhen-
den sensitiven enzymatischen Reaktion nachzuweisen. Durch Markierung mit
GlcDH kann somit auch schnell, billig und effektiv das gewünschte exprimierte
20 Protein nachgewiesen werden.

Darüber hinaus können in einer Reihe von Fällen insbesondere in *E. coli* (GlcDH-
Protein/Polypeptid-Fusionsproteine, , in höherer Ausbeute und Stabilität expri-
miert werden, als ohne GlcDH. Entsprechende Fusionsproteine können somit per
25 se zur Gewinnung und Herstellung von Proteinen / Polypeptiden dienen.

Die *in vivo* Expression von rekombinanten Proteinen spielt eine immer größer
werdende Rolle in der Biotechnologie. Die Fähigkeit, klonierte Genprodukte aus
pro- und eukaryontischen Expressionssystemen wie beispielsweise Bakterien-,
30 Hefe-, Insekten- oder Säugetierzellen, zu erhalten, aufzureinigen und nachzuwei-
sen, wird häufig auch für Studien der Proteinstruktur und -funktion, von Protein-
Protein- und Protein-DNA-Interaktionen sowie Antikörperproduktion und Mutage-

nese verwendet. Mit Hilfe der DNA-Rekombinationstechnik ist es möglich, natürliche Proteine gezielt zu verändern, so daß ihre Funktion verbessert oder variiert wird. Die rekombinanten Proteine werden in ständig weiterentwickelten Expressionssystemen synthetisiert, deren Optimierung an den verschiedensten Stellen im System stattfinden kann.

Der gesamte Prozeß der rekombinanten Proteinsynthese ist in zwei Abschnitte aufteilbar. In einem ersten Schritt findet die molekularbiologische Genisolierung und Expression des Zielproteins statt und im darauffolgenden Schritt der Nachweis und die Aufreinigung aus den rekombinanten Zellen oder ihrem Wachstumsmedium. Auf molekularer Ebene wird das Gen eines Proteins in einen dafür vorgesehenen Expressionsvektor kloniert, anschließend in eine Wirtszelle (Pro- oder Eukaryonten-Zelle) eingeschleust und dort exprimiert. Bakterienzellen erweisen sich dabei als einfache und kostengünstige Systeme, die hohe Ausbeuten liefern. Am häufigsten wird das gramnegative Bakterium *E. coli* als Wirtszelle eingesetzt.

Ziel bei der Expression von Fremdgenen in *E. coli* ist die Gewinnung einer möglichst großen Menge an biologisch aktiven, rekombinanten Proteinen, die sogenannte Überexpression. Bekannt ist, daß eukaryontische Fremdproteine ihre biologische Aktivität dabei durch Aggregation, als Einschußkörperchen, durch inkorrekte Faltung oder proteolytischen Abbau verlieren können. Eine Möglichkeit, diese häufig auftretenden Schwierigkeiten zu vermeiden, bietet die Ausschleusung der exprimierten Proteine als Sekretionsproteine aus der Zelle oder aber die Verwendung sogenannter Fusionsproteine, durch die unlösliche rekombinante Proteine in der Zelle in löslicher Form vorliegen können.

Um die Funktion von Proteinen sowie ihre für die Funktion wichtigen Interaktionspartner zu untersuchen, werden Proteine meistens in eukaryontischen Zellen exprimiert. Dort können die für die Funktion wichtigen posttranskriptionalen Modifikationen und die richtige Kompartimentierung erfolgen. Außerdem sind andere für die korrekte Faltung und Prozessierung wichtige Proteine vorhanden.

Auch bei der Expression von größeren Proteinen und von Proteinen, die posttranskriptionale Modifikationen wie etwa S-S Brückenbildung, Glykosylierung,

Phosphorylierung usw. für die korrekte Faltung benötigen bieten sich eukaryontische Expressionssysteme an. Da diese Systeme in der Regel aufwendig und teuer sind und die Expressionsrate unter der von E.coli liegt, ist es besonders wichtig, ein Nachweissystem zu haben, das schnell, sicher, sensitiv und preiswert ist.

Für den Nachweis von rekombinant gebildeten Fremdproteinen, deren biologische Funktion nicht bekannt ist, existieren zahlreiche Gen-Fusions-Systeme. Darin wird das exprimierte Fusionsprotein über den funktionell bekannten Fusionsproteinanteil nachgewiesen.

Ein empfindliches Nachweissystem ist nötig, um die korrekte Expressierung, die exprimierter Menge, das Molekulargewicht und die funktionelle Aktivität des gebildeten Fusionsproteins zu bestimmen. Die Zahl der funktionell unbekannten Proteine nimmt immer rascher zu und es wird immer wichtiger, dafür schnelle und kostengünstige Nachweissysteme zu entwickeln. Bei den meisten Gen-Fusions-Systemen werden immunologische Verfahren wie z. B. der "enzym-linked-immuno-sorbent-assay" (ELISA) oder der Western-Blot eingesetzt, bei denen rekombinant gebildete Fusionsproteine mit Hilfe von spezifischen Antikörpern detektiert werden.

Entsprechende Fusionsproteine haben aber nicht nur den beschriebenen Vorteil, daß das Fremdprotein indirekt leicht nachgewiesen und analysiert werden kann, sondern sie ermöglichen vielfach, das gewünschte Protein in höheren Ausbeuten zu exprimieren, als es ohne seinen Fusionspartner der Fall wäre. Jeder Fusionspartner hat in einem bestimmten Expressionssystem Vorteile, die er auf den anderen Partner nicht selten zu übertragen in der Lage ist. So kann beispielsweise die Empfindlichkeit mancher Proteine gegenüber proteolytischem Abbau verringert werden, wenn es als Fusionsprotein vorliegt. Auch weisen Fusionsproteine häufig günstigere Löslichkeits- und Sekretionseigenschaften auf als die einzelnen Komponenten.

Es gibt daher zahlreiche Gründe Genfusionen für die Expression rekombinanter Proteine in heterologen Wirten durchzuführen. Diese sind: Erhöhung der Löslich-

keit von Fremdproteinen, Erhöhung der Stabilität von löslichen Fremdproteinen, Lokalisation des Fremdproteins in einem spezifischen Zellabschnitt, schnelle Gewinnung von Fremdproteinen durch vereinfachte Reinigungsstrategien, Möglichkeit der spezifischen Abspaltung des Fusionsproteins, schnelle Nachweismöglichkeit des Fremdproteins aus unaufgereinigten Zellextrakten.

Zur Zeit existieren viele Funktionstests zur Expressionstestung rekombinanter Proteine mit Hilfe von Gen-Fusions-Systemen. Hierbei handelt es sich um einfache Tests, die meist den direkten Nachweis aus ungereinigten Zellextrakten ermöglichen. Die Testsysteme unterscheiden sich jedoch erheblich in Zeitaufwand, Durchsatz und Sensitivität.

Für die oben genannten Zwecke können zwei Arten von Fusionsproteinen unterschieden werden. Zum einen Fusionsproteine, die aus dem gewünschten Protein und einem meist kurzen Oligopeptid bestehen. Dieses Oligopeptid ("Tag") hat die Aufgabe einer Marker- oder Erkennungssequenz für das gewünschte Protein. Zusätzlich kann ein Tag die Reinigung vereinfachen.

Die Hauptanwendung des Tag besteht einmal im Testen der Expression, zum anderen bei der Proteinreinigung. Ein Beispiel hierfür ist der sogenannte His-Tag, der aus einer Peptidsequenz mit sechs aufeinanderfolgenden Histidin-Resten besteht, die direkt mit dem rekombinanten Protein verknüpft ist. Mit Hilfe des angehängten His-Restes kann das Fusionsprotein leicht über eine Metall-Affinitätssäule gereinigt werden (Smith et al., 1988). Eine einfache Detektierung dieses His-Tags findet mit Hilfe des hoch spezifischen monoklonalen Antikörpers His-1 statt (Pogge v. Strandmann et al., 1995). Ein weiterer in Fusionsproteinen verwendeter Marker ist das GFP, ein aus der Qualle *Aequorea victoria* stammendes "green fluorescent protein" (GFP), das als Biolumineszenz-Protein in diversen biotechnologischen Anwendungen eingesetzt wird (Kendall und Badminton, 1998; Chalfie et al., 1994; Inouye et al., 1994). Es kann durch seine Autofluoreszenz einfach detektiert werden in lebenden Zellen, Gelen und sogar lebenden Tieren.

Weitere Beispiele für Tags, die nicht näher erläutert werden sollen, sind das Strep-Tag-System (Uhlén et al., 1990) oder das myc-Epitop-Tag (Pitzurra et al., 1990).

- 5 Die Hauptanwendung von Fusionsproteinen, die aus einem rekombinanten Protein und einem **funktionell aktiven Protein** bestehen, liegt neben dem oben beschriebenen Nachweis in der vereinfachten Reinigung der exprimierten Fusionsproteine. Verschiedene Systeme sind hierunter bekannt, von denen einige nachfolgend kurz erwähnt werden sollen.

10

Im GST-System ermöglichen Fusionsvektoren die Expression von kompletten Genen oder Genfragmenten in Fusion mit Glutathion-S-transferase. Durch Affinitätschromatographie an Glutathion-Sepharose kann das GST-Fusionsprotein leicht aus den Zelllysaten aufgereinigt werden (Smith, Johnson, 1988). Es ist ein
15 biochemischer und ein immunologischer Nachweis verfügbar. Im MBP-System ist das Maltose bindende Protein (maltose-binding protein, MBP) ein periplasmatisches Protein aus *E.coli*, das am Transport von Maltose und Maltodextrinen durch die Bakterienmembran beteiligt ist (Kellermann et al., 1982). Es wurde vor allem für Expression und Reinigung von alkalischer Phosphatase an einer quer-
20 vernetzten Amylose-Säule verwendet. Das Intein-System ist speziell für die schnelle Aufreinigung eines Target-Proteins geeignet. Das Inteingen besitzt die Sequenz für die Intein-Chitin Bindungs Domäne (CBD), wodurch das Fusionsprotein direkt aus dem Zellextrakt an eine Chitinsäule gebunden und damit aufgereinigt werden kann (Chong et al., 1997).

25

Glucose-Dehydrogenase (GlcDH) ist ein Schlüsselenzym während der frühen Phase der Sporenbildung bei *Bacillus megaterium* (Jany et al., 1984). Es katalysiert spezifisch die Oxidation von β -D-Glucose zu D-Gluconolacton, wobei NAD⁺ bzw. NADP⁺ als Coenzym fungieren. Außer in Bakteriensporen, kommt das Enzym
30 auch in der Säugetierleber vor. Es existieren in *B. megaterium* M1286 zwei voneinander unabhängige Glucose-Dehydrogenase-Gene (gdh) (Heilmann et al., 1988). GdhA und gdhB unterscheiden sich in ihrer Nucleotid-Sequenz erheblich, wogegen GlcDH-A und GlcDH-B trotz unterschiedlicher Proteinsequenz annähernd die gleiche Substratspezifität besitzen. Weitere Angaben sowie die ent-

sprechenden DNA- und Aminosäuresequenzen sind auch z. B. der EP-B-0290 768 zu entnehmen.

Die oben beschriebenen Systeme für den Nachweis von rekombinant gebildeten
5 Fremdproteinen, deren biologische Funktion entweder nicht oder nur unzureichend bekannt ist, sind meist kompliziert und zeitaufwendig. Dadurch ist die Verbesserung und Optimierung der Expressionsbedingungen oft nicht schnell oder einfach genug möglich.

10 Deshalb stellt es einen großen Fortschritt dar, einen Fusionsproteinpartner entwickelt zu haben, der einen schnelleren Nachweis des Fusionsproteins ermöglicht, bzw. die im Stand der Technik beschriebenen Nachteile vergleichbarer Systeme nicht aufweist.

15 Es wurde nun gefunden, daß Fusionsproteine, die GlcDH oder eine Sequenz, welche die biologische Aktivität von GlcDH aufweisen, hervorragend geeignet sind, um ein beliebiges gewünschtes "Fremd- oder Zielprotein" rascher, einfacher und damit effizienter nachzuweisen als mit dem beschriebenen Stand der Technik. Diese Eigenschaft beruht auf der überraschenden Erkenntnis, daß GlcDH
20 seine enzymatische Aktivität unter Bedingungen, unter denen andere Enzyme inaktiviert werden (z. B. bei SDS-PAGE), behält.

Bekannt ist die Möglichkeit, Dehydrogenasen aufzureinigen über immobilisierte Farbstoffe wie Cibachron Blue 3 G oder andere NAD analoge Verbindungen wie
25 Aminoethyl-AMP, welche durch ihre Struktur dem Coenzym NAD⁺ ähnlich sind und gleichfalls an alle Dehydrogenasen binden.

Als Teil eines Fusionsproteins erleichtert deshalb Glucose-Dehydrogenase aufgrund seiner Affinität zu diesen z.B. an einem Gel immobilisierten Farbstoffen, welche kommerziell erhältlich sind, die Aufreinigung des Fusionsproteins in einem Schritt. Ferner kann GlcDH als Bestandteil eines Fusionsproteins durch die
30 Kopplung der enzymatischen Reaktion an eine sensitive Farbreaktion, vorzugsweise mit Iodphenylnitrophenyl-phenyltetrazolium-Salz (INT) oder Nitroblau-tetrazolium-Salz (NBT) (unter den genannten Bedingungen) nachgewiesen werden, wodurch sich der indirekte Nachweis des Fremdproteins weiter vereinfacht.

Die Anfärbungsmethode für GlcDH als Markerenzym hat zudem den Vorteil, daß sie die übliche Anfärbung von Proteinen mit z.B. Coomassie-Farbstoffen oder Silberfärbung im selben Gel nicht behindert.

- 5 In einer Ausführungsform der vorliegenden Erfindung besteht das Fusionsprotein neben GlcDH und dem Fremdprotein noch zusätzlich aus einem Tag-Peptid, welches für zusätzliche Charakterisierungen der an das Tag-Peptid gebundenen Proteine verwendet werden kann. Die Charakterisierung erfolgt zum Beispiel über den Polyhistidin-Tag, der von spezifischen Antikörpern als Antigen erkannt wird.
- 10 Der Nachweis des entstandenen Antigen-Antikörper-Komplexes erfolgt dann beispielsweise mit Hilfe einer Peroxidase (POD)-markierten Antikörpers nach an sich bekannten Methoden. Durch die gebundene Peroxidase entsteht nach Zugabe eines entsprechenden Substrats (z.B. ECL-System, Western Exposure Chemiluminescent Detection System, Fa. Amersham) ein chemilumineszierendes
- 15 Produkt, das mit einem hierfür geeigneten Film detektiert werden kann. Der immunologische Nachweis kann aber auch nach an sich bekannter Technik durch ein spezielles Antikörper-Tag, z. B. dem myc-Tag, erfolgen. Der Polyhistidin-Tag, allein oder in Kombination mit dem myc-Tag, hat überdies den Vorteil, daß das Fusionsprotein durch Bindung an einer Metall-Chelatsäule aufgereinigt werden
- 20 kann.

- Das GlcDH-Fusionsprotein kann aber auch direkt an einen spezifischen anti-GlcDH-Antikörper, der z.B. auf einem Chromatographiegel wie Agarose immobilisiert wurde, mittels Affinitätschromatographie aufgereinigt bzw. isoliert werden.

- 25 Ein weiterer Vorteil der Erfindung ist, daß GlcDH vorzugsweise in *E. coli* durch die bekannten Expressionssysteme (s. oben) in hohen Ausbeuten in löslicher Form exprimiert werden kann. So wurde rekombinante Glucose-Dehydrogenase aus *Bacillus megaterium* M1286 erfolgreich in *E. coli* mit hoher enzymatischer
- 30 Aktivität exprimiert (Heilmann 1988). Die Expression anderer eukaryontischer Gene in *E. coli* ist oft durch die Instabilität der Polypeptidkette im bakteriellen Wirt begrenzt. Eine inkorrekte Faltung kann zu Aggregation ("inclusion bodies"), verminderter oder fehlender biologischen Aktivität und proteolytischen Abbau führen. Ein entsprechendes Fusionsgen, bei dem das GlcDH-Gen oder ein Fragment mit

biologischer Aktivität von GlcDH an das Gen des gewünschten Fremdproteins ligiert wurde, kann nun erfindungsgemäß mit nahezu unveränderter Expressionsrate und Ausbeute, verglichen mit dem GlcDH-Gen ohne Fusionspartner, zum Fusionsprotein umgesetzt werden. Dies kann auch dann erfolgen, wenn das
5 Fremdprotein allein an sich nicht oder nur in verminderten Ausbeuten oder nur in inkorrekt gefaltetem Zustand oder nur unter Anwendungen zusätzlicher Techniken exprimiert werden kann. Durch anschließende Abspaltung des Markerproteins GlcDH bzw. des Zielproteins beispielsweise mit Endoproteasen kann somit das gewünschte Fremdprotein erhalten werden.

10

Als Beispiel für ein Zielprotein, welches als Fusionsprotein zusammen mit GlcDH erfolgreich in *E. coli* exprimiert werden kann, dient erfindungsgemäß Tridegin. Tridegin ist ein extrem wirksamer Peptid-Inhibitor für den Blutgerinnungsfaktor XIIIa und stammt aus dem Blutegel *Haementeria ghilianii* (66 AS, 7,6 kD; Finney
15 et al., 1997).

Erfindungsgemäß sind aber keinerlei Einschränkungen in bezug auf die Art und die Eigenschaften des eingesetzten Fremdproteins zu nennen.

20 Die Erfindung beschränkt sich nicht nur auf die Expression der erfindungsgemäßen Fusionsproteine in *E. coli*. Vielmehr können derartige Proteine auch vorteilhaft mittels an sich bekannter Methoden und entsprechender stabiler Vektorkonstruktionen (z. B. mit Hilfe des humanen Cytomegalovirus (CMV)-Promotors) in Säuger-, Hefe- oder Insektenzellen mit guten Expressionsraten synthetisiert werden.
25

Aus dem oben beschriebenen kann die Erfindung demnach wie folgt, bzw. wie in den Patentansprüchen angegeben, zusammenfassend charakterisiert werden:

30 Gegenstand der Erfindung ist somit ein rekombinantes Fusionsprotein, bestehend aus mindestens einer ersten und zweiten Aminosäuresequenz, wobei die erste Sequenz die biologische Aktivität von Glucose-Dehydrogenase aufweist. Gegenstand der Erfindung ist insbesondere ein entsprechendes rekombinantes

Fusionsprotein, bei dem die besagte zweite Sequenz ein beliebiges rekombinantes Protein / Polypeptid X ist oder Teile davon darstellt.

Die erfindungsgemäßen Fusionsproteine können zusätzlich Erkennungssequenzen, insbesondere Tag-Sequenzen enthalten. Gegenstand der Erfindung ist somit ferner ein entsprechendes Fusionsprotein, welches zusätzlich mindestens eine weitere für die Detektion geeignete Erkennungssequenz oder Tag-Sequenz aufweisen kann.

Die erfindungsgemäßen Fusionsproteine können verschiedenartig eingesetzt werden. Dabei spielt die Glucose-Dehydrogenase mit ihren Eigenschaften die entscheidende Rolle. So ist Gegenstand der Erfindung die Verwendung von Glucose-Dehydrogenase als Detektorprotein für ein beliebiges rekombinantes Protein / Polypeptid X in einem der besagten Fusionsproteine. Weiter ist Gegenstand der Erfindung die Verwendung von Glucose-Dehydrogenase in einem Nachweissystem für die Expression eines rekombinanten Proteins / Polypeptids X als Bestandteil eines entsprechenden Fusionsproteins. Ferner ist Gegenstand der Erfindung die Verwendung von GlcDH zum Nachweis für Protein-Protein Interaktionen, wobei ein Partner dem rekombinanten Protein / Polypeptid X, wie oben und unten definiert, entspricht. Schließlich kann GlcDH entsprechend der Erfindung als Detektorprotein für ein beliebiges drittes Protein / Polypeptid dienen, welches nicht Bestandteil des Fusionsproteins ist, aber an die zweite Sequenz des Proteins / Polypeptids X des besagten Fusionsproteins zu binden vermag. Ferner kann GlcDH als Markerprotein von einem Partner in ELISA-Systemen, Western-blot und verwandten Systemen eingesetzt werden.

Die Erfindung umfaßt, da sie rekombinante Techniken einsetzt, natürlich auch entsprechende Vektoren, Wirtszellen und Expressionssysteme. Gegenstand der Erfindung ist neben diesen Vektoren und Wirtszellen als solche auch die Verwendung entsprechender Expressionsvektoren bei der Optimierung der Expression eines rekombinanten Proteins / Polypeptids X in einem rekombinanten Herstellungsverfahren sowie die Verwendung einer entsprechenden Wirtszelle bei der Optimierung der Expression eines rekombinanten Proteins / Polypeptids X in einem solchen Herstellungsverfahren.

Gegenstand der Erfindung ist auch ein Verfahren zum schnellen Nachweis eines beliebigen rekombinanten Proteins / Polypeptids X mittels Gelelektrophorese, insbesondere SDS-PAGE-Gelelektrophorese, wobei ein entsprechendes Fusionsprotein hergestellt, mittels Gelelektrophorese aufgetrennt wird und das nachzuweisende rekombinante Protein / Polypeptid im Gel über die Enzymaktivität der Glucose-Dehydrogenase sichtbar gemacht wird.

Erfindungsgemäß wird dabei zum Nachweis der Enzymaktivität der Glucose-Dehydrogenase eine Farbreaktion auf Basis von Tetrazoliumsalzen, insbesondere Iodphenylnitrophenyl-phenyltetrazolium-Salz (INT) oder Nitroblau-tetrazolium-Salz (NBT), eingesetzt, wobei sich gegebenenfalls vor oder nach der besagten erfolgten Farbreaktion eine generelle Proteinfärbung gemäß des Standes der Technik anschließen kann.

Im folgenden sind die Abbildungen kurz erklärt:

Abb. 1: Konstruktionsschema des Vektors pAW2. Der Vektor enthält die Sequenz für GlcDH. Die vollständige Sequenz ist in Abb. 9 dargestellt.

Abb. 2: Konstruktionsschema des Vektors pAW3.

Abb. 3: Konstruktionsschema des Vektors pAW4. Der Vektor enthält die Sequenz für GlcDH und Tridegin. Die vollständige Sequenz ist in Abb. 10 dargestellt.

Abb. 4: Anfärbung von GlcDH auf einem SDS-PAA-Gel. Die Färbemethode ist in den Beispielen näher beschrieben. 1: Rainbow-Marker; 2: 0,1 µg GlcDH; 3: 0,05 µg GlcDH; 4: 0,001 µg GlcDH; 5: Lysat HC11-Zellen; 6: prestained SDS-Marker.

Abb. 5: Nachweis des exprimierten GlcDH-Enzyms (15% SDS-PAA-Gel, INT-Färbung); 1: Rainbow-Marker; 2: 0,2 µg native GlcDH; 3: 10µl Zellextrakt /

1ml Suspension Klon 2, 4: 10µl Zellextrakt / 1ml Suspension Klon1; 5 : pre-stained SDS-Marker; Zellextraktvolumen: 100 µl.

Abb. 6: Verdünnungsreihe aus pAW2-Expression (15% SDS-PAA-Gel, INT-Färbung); 1: Rainbow-Marker; 2: 10 µl Zellextrakt / 100 µl Suspension; 3: 10 µl Zellextrakt / 1:5 Verdünnung; 4: 10 µl Zellextrakt / 1:10 Verdünnung; 5: 10 µl Zellextrakt / 1:20 Verdünnung; 6: 0,5 µg GlcDH; 7: Broad-Range SDS-Marker; 8: prestained SDS-Marker; Zellextraktvolumen: 100 µl.

Abb. 7: Nachweis des exprimierten Tridegin/GlcDH-Fusionproteins (10% SDS-PAA-Gel, INT/CBB); 1: Broad-Range SDS-Marker; 2: 1 µg GlcDH; 3: 0,5 µg GlcDH; 4: 0,1 µg GlcDH; 5: 500 µl Zellextrakt; 6: 200 µl Zellextrakt; 7: 100 µl Zellextrakt; 8: 500 µl Zellextrakt (pAW2-Expression); Zellextraktvolumen: 100 µl.

Abb. 8: Immundetektion von Tridegin/His- und Tridegin/His/GlcDH-Fusionsprotein (aus 10% SDS-PAA-Gel, ECL-Detektion) und Vergleich mit Tridegin/His/GlcDH (10% SDS-PAA-Gel, INT-CBB-Färbung); 1: Broad-Range-Marker; 2: 1 ml Zellextrakt (pAW2-Expression); 3: 100 µl Zellextrakt (pST106-Expression); 4: 200 µl Zellextrakt (pST106-Expression); 5: 300 µl Zellextrakt (pAW4-Expression); 6: 2,5 µg Calin-His-Positivkontrolle; 7: Broad-Range-Marker; 8: 100 µl (pAW4-Expression); Zellextraktvolumen: 100 µl.

Abb. 9: Sequenzbeschreibung von Plasmid pAW2.

Abb. 10: Sequenzbeschreibung von Plasmid pAW4.

Im folgenden sind die oben und unten verwendeten Abkürzungen erläutert

A	Adenin
A _x	Absorption bei x nm
AK	Antikörper
Amp	Ampicillin
AP	Alkalische Phosphatase
APS	Ammoniumperoxodisulfat

	AS	Aminosäure
	bla	β -Lactamase-Gen
	BIS	N,N'-Methylenbisacrylamid
	bp	Basenpaare
5	BSA	Bovines (Rinder)-Serumalbumin
	C	Cytosin
	cDNA	copy (complementary) DNA
	CBB	Coomassie Brilliant Blue
	CIP	calf intestinal phosphatase
10	dNTP	2'-Desoxyribonucleosid-5'-triphosphat
	ddNTP	2',3'-Desoxyribonucleosid-5'-triphosphat
	DMF	Dimethylformamid
	DMSO	Dimethyl-Sulfoxid
	DNA	Desoxyribonucleinsäure
15	dsDNA	Doppelstrang-DNA
	DTT	Dithiothreitol
	ECL	Exposure TM Chemiluminescence
	EDTA	Ethylendiamin-N,N,N',N'-Tetraessigsäure, Dinatriumsalz
	ELISA	Enzyme linked immuno sorbent assay
20	EtBr	Ethidiumbromid
	EtOH	Ethanol
	f. c.	final concentration
	FACS	Fluorescent activated cell sorting
	G	Guanin
25	GFP	Green fluorescent protein
	GlcDH	Glucose-Dehydrogenase (Protein)
	gdh	Glucose-Dehydrogenase (Gen)
	GST	Glutathion-S-Transferase
	His	Histidin-Rest
30	HRP	Horseradish peroxidase
	IB	Inclusion Body
	IgG	Immunglobulin G
	INT	Iodnitrotetrazoliumviolett
	kb	Kilobasenpaare

	kD	KiloDalton
	mA	Milliampere
	m-RNA	messenger-RNA
	MBP	Maltose-binding protein
5	MCS	Multiple cloning site
	M _r	relatives Molekulargewicht
	NAD (P)	Nicotinamid-adenin-dinucleotid (-Phosphat), freie Säure
	Od _x	optische Dichte bei x nm
	ompA	outer membrane protein A
10	ori	origin of replication
	PAA	Polyacrylamid
	PAGE	Polyacrylamid-Gelelektrophorese
	PCR	Polymerase Chain Reaction
	POD	Peroxidase
15	PVDF	Polyvinylidendifluorid
	RNA	Ribonucleinsäure
	RNAse	Ribonuclease
	rpm	Umdrehungen pro Minute
	rRNA	ribosomale RNA
20	RT	Raumtemperatur
	SDS	Natriumdodecylsulfat
	ssDNA	Einzelstrang-DNA
	Strep	Streptavidin
	T	Thymin
25	T _m	Schmelzpunkt (DNA-Duplex)
	t-RNA	transfer-RNA
	Taq	<i>Thermophilus aquaticus</i>
	TCA	Trichloressigsäure
	TEMED	N,N,N',N'-Tetramethylethylendiamin
30	Tet	Tetracyclin
	Tris	Tris-(hydroxymethyl)-aminomethan
	U	Unit = Einheit der Enzymaktivität
	U	Uracil
	UV	Ultraviolette Strahlung

ÜN	Über Nacht
V	Volt
VIS	visible (sichtbarer Bereich)
w/v	weight per volume

5

Literaturverzeichnis:

- Aoki et al. (1996), FEBS Letters **384**, 193-197
- Banauch et al. (1975), Z. Klin. Chem. Klin. Biochem. **13. Jg.**, 101-107
- Bertram, & Gassen (1991) Gentechnische Methoden,
10 Eine Sammlung von Arbeitsanleitungen für das molekularbiologische Labor. Gustav Fischer Verlag, Stuttgart, Jena, New York
- Brewer & Sassenfeld (1985), Trends in Biotechnology **3**, No. 5, 119-122
- Brown, T. A. (1993) Gentechnologie für Einsteiger: Grundlagen, Methoden, Anwendungen. Spektrum Akademischer Verlag, Heidelberg; Berlin; Oxford
- 15 Casadaban et al. (1990), Methods in Enzymology **100**, 293
- Chalfie et al. (1994), Science **263**, 802-805
- Chong, S. et al. (1997), Gene **192**, 271-281
- Collins-Racie et al. (1995), Biotechnology **13**, 982-987
- Di Guan et al. (1988), Gene **67**, 21-30
- 20 Ettinger et al. (1996), Proc. Natl. Acad. Sci. USA **93**, 13102-13107
- Finney et al. (1997), Biochem. J. **324**, 797-805
- Gazitt et al. (1992), Journal of Immunological Methods **148**, 159-169
- Ghosh et al. (1995), Analytical Biochemistry **225**, 376-378
- Goeddel et al. (1979), Proc. Natl. Acad. Sci. U.S.A. **76**, 106-110
- 25 Hafner & Hoff (1984), Genetik. Neubearbeitung, Schrödel-Verlag, Hannover
- Harlow & Lane (1988), A Laboratory Manual, Cold Spring Harbor
- Harris & Angal (1990) Protein purification applications: a practical approach. Oxford University Press, Oxford; New York; Tokyo
- Heilmann et al. (1988), Eur. J. Biochem. **174**, 485-490
- 30 Hilt et al. (1991), Biochimica et Biophysica Acta **1076**, 298-304
- Ibelgauf, H. (1990) Gentechnologie von A bis Z. Erweiterte Ausgabe, VCH-Verlag, Weinheim
- Inouye et al. (1994), FEBS Letters **341**, 277-280
- Itakura et al. (1977). Science **198**, 1056-1063

- Jany et al. (1984), FEBS Letters **165**, no. 1, 6-10
- Kellermann & Ferenci (1982), Methods in Enzymology **90**, 459-463
- Laemmli (1970), Nature **227**, 680-685
- La Vallie & McCoy, (1995), Curr. Opin. Biotechnol. **6**, 501-506
- 5 Makino et al. (1989), Journal of Biological Chemistry **264**, No. 11, 6381-6385
- Marston (1986), Biochem. J. **240**, 1-12
- Moks et al. (1987), Biochemistry **26**, 5239-5244
- Okorokov et al. (1995), Protein Expr. Purif., **6**, 472-480
- Pharmacia Biotech 1: From Cells to Sequences, A Guide to PCR analysis
10 of nucleic acids
- Pharmacia Biotech 2: The Recombinant Protein Handbook, Principles and Methods
- Pitzurra et al. (1990), Journal of Immunological Methods **135**, 71-75
- Pogge v. Strandmann et al. (1995), Protein Engineering **8**, No.7, 733-735
- 15 Sambrook et al., (1989) Molecular Cloning, A Laboratory Manual 1,
Second Edition, Cold Spring Harbor Laboratory Press, USA
- Sanger et al. (1977), Proc. Natl. Acad. Sci. U. S. A. **74**, 5463-5467
- Schein (1989), Bio/Technology **7**, 1141-1149
- Scopes (1994) Protein purification: principles and practice, 3rd ed.,
20 Springer-Verlag, New York; Berlin; Heidelberg
- Smith & Johnson (1988), Gene **67**, 31-40
- Smith et al. (1988), Journal of Biological Chemistry **263**, No. 15, 7211-7215
- Uhlén & Moks (1990), Gene Fusions for Purpose of Expression:
An Introduction [12]. Methods in Enzymology **185**, 129-143
- 25 Uhlén et al. (1983), Gene **23**, 369

Falls nicht anderweitig dargelegt, entsprechen die bei dieser Erfindung verwendeten Methoden und Techniken hinlänglich bekannten und in der einschlägigen Literatur beschriebenen Methoden und Verfahren. Insbesondere sind die Offen-

30 barungsgehalte der oben genannten Publikationen und Patentanmeldungen, vor allem von Sambrook et al. und Harlow & Lane sowie die EP-B-0290 768 erfindungsgemäß mit umfaßt. Die gemäß der Erfindung verwendeten Plasmide und Wirtszellen sind in der Regel exemplarisch und können im Prinzip durch modifizierte oder anders aufgebaute Vektorkonstruktionen oder andere Wirtszellen er-

setzt werden, sofern sie noch die genannten erfindungswesentlichen Bestandteile aufweisen. Die Herstellung solcher Vektorkonstruktionen sowie die Transfektion entsprechender Wirtszellen und die Expression und Aufreinigung der gewünschten Proteine entsprechen weitgehend bekannten Standardtechniken und können
5 erfindungsgemäß ebenfalls innerhalb eines weiten Rahmens modifiziert werden.

Im folgenden wird die Erfindung näher beschrieben. Weitere Details sind in den Beispielen erläutert.

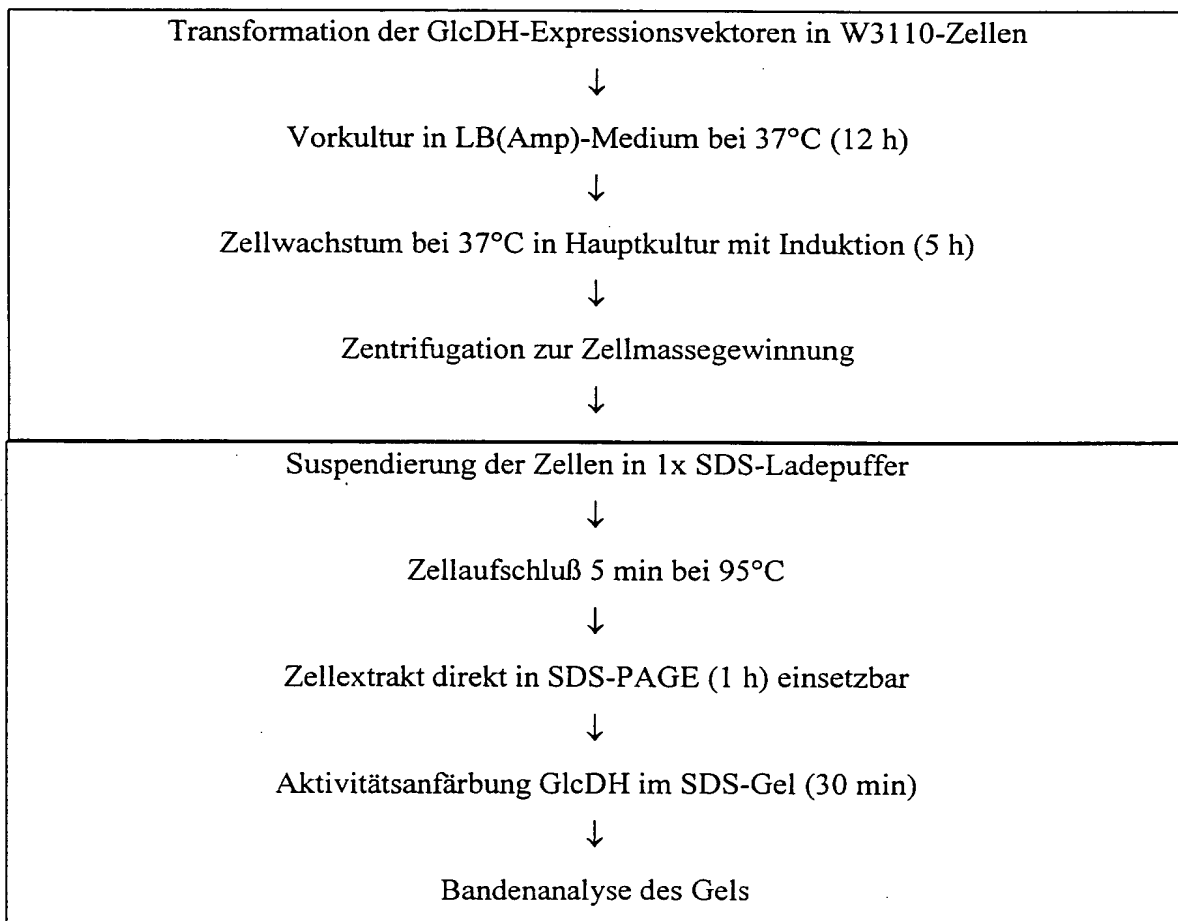
10 Das *Bacillus megaterium* GlcDH-Strukturgen wurde mittels PCR modifiziert, wobei das Plasmid pJH115 (EP 0290 768) als Template fungierte. Das amplifizierte Fragment (0,8 kb), das an einem Ende eine PstI- und am anderen eine Eco47III-Erkennungssequenz besaß, wurde mit diesen Enzymen verdaut und in den cytoplasmatischen (pRG45) oder periplasmatischen (pST84) *E. coli*-
15 Expressionsvektor kloniert (Abb 1, 2). Die resultierenden Plasmide, pAW2 und pAW3, besaßen nun ein GlcDH-Gen, das ein Protein von etwa 30 kD (261 AS) verschlüsselt und unterhalb des starken Tet-Promotors liegt. Der cytoplasmatische pAW2-Expressionsvektor besitzt eine Größe von ca. 4 kb. Der periplasmatische pAW3-Sekretionsvektor ist geringfügig größer und unterscheidet sich von
20 pAW2 nur in einer der multiple-cloning-site (MCS) vorgeschalteten omp A-Signalsequenz, die dazu führt, daß das rekombinante Protein in das Periplasma sezerniert werden kann. Beide Vektoren besitzen zudem eine MCS mit 12 unterschiedlichen Restriktionsschnittstellen, die in frame-Klonierung mit dem nachfolgenden His-Tag ermöglichen. Durch den Polyhistidin (6His)-Tag wird die Reini-
25 gung des rekombinanten Proteins an einer Metall-Affinitätssäule möglich. Der Vektor pAW4 enthält schließlich das Tridegin-Gen und das GlcDH-Gen, welche über eine MCS miteinander verbunden wurden sowie den Polyhistidin (6 His)-Tag, der stromabwärts mit dem GlcDH-Gen ligiert ist. Die einzelnen Konstruktionen sind in den Abb. 1, 2 und 3 dargestellt. Die gewählten Plasmidkonstruktionen
30 sind aber nur beispielhaft und schränken die Erfindung nicht ein. Sie können durch andere geeignete Konstruktionen, welche die genannten DNA-Sequenzen enthalten, ersetzt werden. Die Herstellung der Vektoren, der Klone und die Expression der Proteine ist in den Beispielen weiter spezifiziert.

Die Empfindlichkeit der Aktivitätsanfärbung wurde im reduzierten SDS-Gel für native GlcDH durchgeführt. Hierfür wurde eine Konzentrationsreihe mit der nativen GlcDH ($c = 1 \text{ mg/ml}$; $A = 200 \text{ U/ml}$) angefertigt und eine Negativkontrolle vorbereitet. Nach SDS-PAGE und Aktivitätsanfärbung mittels INT wurde das in Abb. 3 dargestellte SDS-Gel erhalten. Mit Hilfe des eingesetzten Tests konnte die GlcDH bis zu einer Konzentration von 50 ng nachgewiesen werden. Die Negativkontrolle, in der keine GlcDH vorhanden ist, zeigt wie erwartet keine Bande auf.

Mit Hilfe von Markerproteinen läßt sich über eine Eichkurve das genaue Molekulargewicht der nativen GlcDH ermitteln. Hierfür wurden die relativen Laufstrecken der Markerproteine bestimmt und gegen deren zugehöriges logarithmisches Molekulargewicht aufgetragen.

Die durchgeführten Expressionen wurden gemäß dargestellten Schema ausgeführt (Tab. 1):

Tab. 1



Das Plasmid pAW2/Klon9 (pAW2/K9) wurde in den kompetenten *E. coli*-Expressions-stamm W3110 transformiert und zwei Klone von der erhaltenen Transformationsplatte zur Beimpfung einer 5 ml-Vorkultur verwendet. Die Anhydrotetracyclin-Induktion fand 2 h nach der Beimpfung der Hauptkultur statt. Die gesamte Expression dauerte 5 h und wurde bei einem OD-Wert von 1,65 für Klon 1 und 1,63 für Klon 2 abgebrochen. Nach SDS-PAGE und GlcDH-Aktivitätsanfärbung konnten je Klon eine starke GlcDH-Bande (ca. 35 kD) aus 1 ml Zellsuspension nachgewiesen werden.

Bei der Durchführung von SDS-PAGE unter reduzierten und nicht reduzierten Bedingungen wurde kein Unterschied zwischen den erhaltenen GlcDH-Banden deutlich. Hierfür wurden jeweils 500 bis 100 µl der Zellsuspension im SDS-Gel durch GlcDH-Aktivitätsanfärbung mit INT untersucht.

Um die Empfindlichkeit der GlcDH-Aktivitätsanfärbung gegenüber der Coomassie-Färbung zu verdeutlichen wurden Proben aus 100 µl Zellsuspension, sowie 1/5-, 1/10- und 1/20-Verdünnungen der Zellsuspension hergestellt. Das Endvolumen der Verdünnungen betrug ebenfalls 100 µl. Mit dem erhaltenen SDS-Gel wurde nach der GlcDH-Aktivitätsanfärbung eine Coomassie-Färbung durchgeführt, um weitere Proteinbanden sichtbar zu machen. Das daraus resultierende SDS-Gel ist in Abbildung 4 dargestellt. Mittels der GlcDH-Aktivitätsanfärbung ist bei der 1/20-Verdünnung noch eine deutliche Bande erkennbar, wohingegen Coomassie-gefärbte Banden kaum noch wahrnehmbar sind.

Das *Haementeria ghilianii* Tridegin-Strukturgen mit gekoppeltem His-Tag wurde mittels PCR modifiziert, wobei das Plasmid pST106 als Template fungierte. Das amplifizierte Fragment (0,25 kb), das von einer ClaI- und PstI-Erkennungssequenz flankiert wird, wurde mit diesen Enzymen verdaut und in den cytoplasmatischen *E. coli*-GlcDH-Fusionsvektor pAW2 kloniert. Das resultierende Plasmid pAW4 besaß nun ein Tridegin-His-GlcDH-Fusionsproteingen, das für ein Protein von etwa 44 kD codiert und unterhalb des starken Tet-Promotors liegt. Der Zellextrakt aus dem *E. coli*-Stamm W 3110, der das cytoplasmatische pAW4-Plasmid beinhaltet, wurde mit Hilfe der SDS-PAGE und GlcDH-Aktivitätsanfärbung analysiert. Dabei konnten mehrere rotviolett gefärbte Banden bei 35, 37, 40 und 43 kD nachgewiesen werden. Bei der 43 kD-Bande handelte

es sich um das gewünschte Tridegin-His-GlcDH-Fusionsprotein, dessen Molekulargewicht jedoch etwas kleiner als der theoretische Wert von 44 kD war. Die restlichen nachweisbaren Banden wurden vermutlich durch einen proteolytischen Abbau des Fusionsproteins im *E. coli* erzeugt, denn die kleinste angefärbte Bande von 35 kD entspricht in etwa der Größe der GlcDH. Aufgrund eines Größenvergleichs konnte die gebildete 37 kD-Bande als das His-GlcDH-Abbauprodukt identifiziert werden.

Die Durchführung einer Expressionskinetik ergab, daß 2 Stunden nach der Induktion des Tet-Promotors mit Anhydrotetracyclin der proteolytische Abbau des gebildeten Fusionsproteins eintrat, d.h. ab diesem Zeitpunkt waren zusätzliche Banden im SDS-Gel durch Aktivitätsanfärbung nachweisbar. Das gebildete Fusionsprotein war nicht gegenüber den *E. coli*-Proteasen stabil, was sich in seinem relativ schnellen Proteinabbau zeigt. Durch den Einsatz des konstruierten periplasmatischen GlcDH-Fusionsvektors pAW3, konnte der proteolytische Abbau des Fusionsproteins in der Zelle vermieden werden, da hierbei das exprimierte Fusionsprotein in den periplasmatischen Zwischenraum der *E. coli*-Zellen sezerniert würde. Die *E. coli*-Proteasen befinden sich vornehmlich im Cytoplasma.

Die Sensitivität und Spezifität des GlcDH-Fusionsprotein-Nachweises ermöglichen ein schnelles und einfaches Screening von rekombinanten Fremdproteinen. Die Sensitivität des GlcDH-Nachweissystems wurde mit Hilfe von nativer GlcDH bestimmt. Der Aktivitätsnachweis der nativen GlcDH ergab im SDS-PAA-Gel eine rotviolett gefärbte Bande bei ca. 30-35 kD.

Die cytoplasmatische Expression im *E. coli*-Stamm W3110 der rekombinanten GlcDH aus pAW2 ergab das gleiche Molekulargewicht. Der Sensitivitätsvergleich der nativen GlcDH zur rekombinanten GlcDH konnte durch einen Vergleich der Bandenintensitäten stattfinden.

Das entwickelte Testsystem (siehe Beispiele) bietet zudem die Möglichkeit, eine doppelte Anfärbung der SDS-Gele durchzuführen. In der ersten Färbung findet die spezifische Detektion der GlcDH-Banden statt. Zur Hintergrundfärbung kann anschließend eine übliche Proteinfärbung, z. B. eine Coomassie-Anfärbung der restlichen Proteine stattfinden. Die GlcDH behält unter reduzierenden Bedingungen in Gegenwart von SDS überraschenderweise und erfindungsgemäß ihre

vollständige Aktivität, wodurch der schnelle Nachweis im SDS-Gel ermöglicht wird.

- Erfindungsgemäß ist es weiterhin möglich, die Empfindlichkeit des GlcDH-Aktivitätsnachweises, durch Verwendung von Nitroblautetrazolium-Salz (NBT) als Substrat für die GlcDH, zu steigern. Die Reaktionsgeschwindigkeit des GlcDH-Nachweises mittels INT kann jedoch durch den Einsatz von Triton X-100 (1% Endlösung) oder NaCl-Zugabe (1 M Endlösung) weiter gesteigert werden.
- 10 Die rekombinanten Fusionsproteine Tridegin/His und Tridegin/His/GlcDH wurden durch Expression des pST106- und pAW4-Plasmids gewonnen (Abb. 1, 2). Nach Zellaufschluß des jeweiligen Expressionsansatzes wurden die Proben in der SDS-PAGE aufgetrennt und auf eine Membran transferiert. Das Tridegin-His-GlcDH-Fusionsprotein konnte über seinen enthaltenen His-Tag immunologisch
- 15 durch die Verwendung eines Anti-^{RGS}His-Antikörpers im Western-Blot nachgewiesen werden. Als Kontrollen dienten gereinigtes rekombinantes Calin (Blutegelprotein), welches einen terminalen His-Tag besitzt, sowie der Zellextrakt der exprimierter rekombinanter GlcDH, die keinen His-Tag besitzt. Der Anti-^{RGS}His-Antikörper konnte für das rekombinante Tridegin/His/GlcDH-Fusionsprotein eine
- 20 Bande bei ca. 37 kD und eine weitere Bande bei ca. 43 kD detektieren (Abb. 6). Vergleicht man die erhaltenen Bandengrößen mit den nach Aktivitätsanfärbung im SDS-Gel erhaltenen Banden, so zeigt sich, daß die 43 kD-Bande das Tridegin-His-GlcDH-Fusionsprotein und die 37 kD-Bande das His-GlcDH-Abbauprodukt des gesamten Fusionsproteins darstellt. Das Calin/His-Tag-Protein
- 25 ergab eine Bande mit ca. 26 kD. Das etwas kleinere rekombinante Tridegin/His-Tag-Protein ergab eine Bande mit ca. 23 kD, sowie weitere Banden, die auf eine Bindung des His-Antikörpers mit weiteren exprimierten Proteinen hinweisen. Der immunologische Nachweis mit dem Anti-^{RGS}His-Antikörper beweist also, daß das bei 43kD und das bei 37 kD detektierte Protein einen His-Tag enthält. Zu-
- 30 dem entsprach diese Proteingröße annähernd der theoretischen Größe (36,5 kD) des GlcDH-Proteins mit gekoppeltem His-Tag.

Zusätzlich zum Expressions-Nachweis des rekombinanten Tridegins wurde die biologische Aktivität des Tridegins als Bestandteil des Tridegin-GlcDH Fusionsproteins, im speziellen Fall aus pAW4, untersucht. Dieser Test beruht auf der Hemmung von Faktor XIIIa durch natives Drüsenhomogenat aus Blutegeln, bzw.
5 gereinigtes Tridegin (Finney et al., 1997). Der modifizierte Test ist in den Beispielen beschrieben. Zur Kontrolle wurde das entsprechende Fusionsprotein aus pST106 und das GlcDH-Protein aus pAW2 exprimiert. Im Vergleich der enzymatischen Aktivität mit rekombinanten Tridegin, das entweder als GlcDH-Tridegin Fusionsprotein oder als Tridegin-His-Tag in *E. coli* exprimiert worden ist., konnten
10 keine wesentlichen Unterschiede festgestellt werden. Darüberhinaus zeigten die rekombinanten Tridegin-Proteine aus den beiden unterschiedlichen Expressionen vergleichbare biologische Aktivitäten wie das native Homogenat aus Blutegeldrüsen. Daraus kann gefolgert werden, daß die Fusion mit GlcDH keinerlei störenden Einfluß auf die biologische Aktivität des coexprimierten Fremdgens ausübt.

15 Tridegin selbst (d. h. nicht als Fusionsprotein) besitzt nach Durchführung einer *E. coli*-Expression keine Aktivität und wird als Inclusion Body gebildet. Wird GlcDH in *E. coli* exprimiert, erhält man ein Enzym mit hoher spezifischer Aktivität und Stabilität in löslicher Form. In Expressionsversuchen konnte nachgewiesen werden, daß Proteine, die ein hohes Löslichkeitsvermögen bei der *E. coli*-Expression besitzen, das Löslichkeitsvermögen der Fremdproteinexpression erhöhen, wenn
20 sie mit diesen fusioniert werden (LaVallie, 1995). Die Fusion von Tridegin mit GlcDH erhöhte auch in diesem Fall die Löslichkeit des Tridegins, denn durch einen biologischen Nachweis, bei dem Tridegin den Faktor XIIIa inhibiert, konnte
25 die Aktivität des Tridegins nach der *E. coli*-Expression als Tridegin-His-GlcDH-Fusionprotein nachgewiesen werden. Das GlcDH-Fusionsprotein wird in *E. coli* in hoher Ausbeute exprimiert.

Die Möglichkeit, klonierte Gene als Fusionsproteine zu exprimieren, die ein Protein von bekannter Größe und biologischer Funktion enthalten, vereinfacht den
30 Nachweis des Genprodukts merklich. Aus diesem Grund sind, wie einleitend bereits erwähnt, zahlreiche Fusionsexpressions-Systeme entwickelt worden, die verschiedene Nachweisstrategien aufweisen.

Im Vergleich zu den bekannten Systemen stellt sich das erfindungsgemäße GlcDH-Fusionssystem in E. coli wie in Tab. 2 gezeigt, dar. In einigen Systemen kann das N-terminale Fusionsprotein vom C-terminalen Ziel- oder Fremdprotein abgespalten werden (Collins-Racie et al., 1995).

5 Tab 2:

Tag/Fusions-partner	MW (kD)	Nachweis	Vorteil
GlcDH	30	Funktionstest im SDS-Gel	Schnell und billig, direkter Nachweis im SDS-Gel
His-Tag (Pogge v. Strandmann et al., 1995)	1-7	Western-blot, ELISA	klein
Strep-Tag (Uhlén et al., 1990)	13	Western-blot	klein
myc-Epitop (Pitzurra et al., 1990; Gazitt et al., 1992)	1-2	Western-blot, ELISA	klein
IgG-Teile, Fc (Moks et al., 1987; Ettlinger et al., 1996)	2-5	Western-blot, ELISA	klein, Selektion von Zellen (FACS)
GFP (Chalfie et al., 1994; Inouye et al., 1994)	27	Fluoreszenz, Western-blot	Selektion von Zellen bereits in Kulturschale, mehrere gleichzeitig nachweisbar (FACS)
Intein (Chong et al., 1997)	48	Western-blot	Fusionspartner kann entfernt werden
GST (Smith, Johnson, 1988; Gosh et al., 1995)	26	Western-blot, colorimetrischer Nachweis in Lösung	Fusionspartner kann entfernt werden
MBP (Chu di Guan et al., 1988; Kellermann et al., 1982)	40	Western-blot	Fusionspartner kann entfernt werden

Verfahren	Voraussetzung	Zeitbedarf	Durchsatz	Sensitivität	Information
GlcDH-Nachweis	GlcDH funktionell aktiv	ca. 3 h	Mittel - Hoch	50 ng	Proteinmenge + Proteingröße
ELISA	2 Antikörper	ca. 1 Tag	Hoch	pg-ng	Proteinmenge
Western-blot	1-2 Antikörper Tag am Protein	1-2 Tage	Klein	ng	Proteingröße + Proteinmenge

Ein sehr großer Vorteil des erfindungsgemäßen GlcDH-Nachweissystem ist die Tatsache, daß hierfür wie z.B. für den Nachweis mittels Western-blot keine Antikörper oder sonstige Materialien wie z.B. Membranen, Blot-Apparatur, Entwicklungs-
5 lermaschine mit Filmen, Mikrotiterplatten, Titerplattenlesegerät, u.s.w. benötigt werden. Dadurch entwickelt sich der Nachweis rekombinanter Fusionsproteine mit dem GlcDH-System sehr viel günstiger und schneller. Mit Hilfe des GlcDH-Nachweises kann neben der Information über die Menge des exprimierten Fusionsproteins, auch die entsprechende Größe des Fusionsproteins direkt im SDS-
10 PAA Gel ohne Transfer auf eine Membran festgestellt werden. Ist die Aktivität der GlcDH im Fusionsprotein nachweisbar, so sollte der Fusionspartner in der Regel auch funktionell aktiv sein. GlcDH stört die Faltung des Fusionspartners nicht. Nachfolgend (Tab. 3, unten) wurde ein effizientes Verfahren zur Gewinnung und zum Nachweis eines in *E. coli* gewonnenen Fusionsproteins aus der Literatur
15 ausgewählt, welches die Vorteile des erfindungsgemäßen GlcDH-Fusionsproteinsystems in einem Vergleich zeigt.

Das erfindungsgemäße GlcDH-Fusionsproteinsystem ist ferner besonders geeignet, die Löslichkeit von Proteinen zu erhöhen, die insbesondere in *E. coli* als In-
20 clusion Bodys gebildet werden und deshalb eine anschließende Proteinaufreinigung schwierig und teuer gestalten. Normalerweise müssen Proteine, die als Inclusion Bodys gebildet wurden, durch aufwendige Verfahren in ihren nativen Zustand überführt werden. Dies entfällt bei Anwendung der Fusionsproteine gemäß der Erfindung.

25

Zusammenfassend stellen sich die Vorteile der erfindungsgemäßen Fusionsproteine in ihrer Verwendung als GlcDH-Nachweissystem wie folgt, dar.

- Stabilität unter SDS und reduzierenden (denaturierenden) Bedingungen
- Sensitiver GlcDH-spezifischer enzymatischer Farbstest
- 30 • Sensitivität bis mindestens 50 ng
- Schneller Nachweis direkt im SDS-Gel mit Bestimmung des Molekulargewichts des Fusionspartners
- Möglichkeit zusätzlicher Proteinfärbungen

- Kostengünstige Materialien, geringer apparativer Aufwand
- gute Expression in E. coli, einschließlich des Zielproteins unter Erhaltung der biologischen Aktivität
- Möglichkeit der Vermeidung von Inclusion Bodies des Fremd-/Zielproteins oder anderer durch falsche Faltung erzeugte Aggregate.
- Möglichkeit der Aufreinigung des Fusionsproteins über Affinitätschromatographie z.B. an Farbstoffen (Cibacron Blue 3G)

Tab. 3

Konstruktion/Transformation des <u>ProteinA/GFP-Fusionsvektors</u>	Konstruktion/Transformation des <u>GlcDH/Tridegin-Fusionsvektors</u>
↓	↓
Wachstum der Zellen auf LB-Agarplatten bei 37°C (1 Tag)	Vorkultur in LB(Amp)-Medium bei 37°C (12 h)
↓	↓
Zellwachstum bei 25°C (3 Tage)	Zellwachstum bei 37°C in Hauptkultur mit Induktion (5 h)
↓	↓
Suspendierung der Zellen in Puffer (pH 8,0)	Suspendierung der Zellen in 1x SDS-Ladepuffer
↓	↓
Zellaufschluß und Abzentrifugation der Zelltrümmer	SDS-Zellaufschluß 5 min bei 95°C
↓	↓
SDS-PAGE zur Proteintrennung (1 h)	SDS-PAGE (1 h) mit Zellextrakt
↓	↓
Proteintransfer auf Nitrocellulose-Membran (1 h)	Aktivitätsanfärbung GlcDH im SDS-Gel (30 min)
↓	↓
Blockierungsreaktion (1 h)	
↓	
Antikörperreaktion (1 h)	
↓	
Inkubation in ProteinA-GFP-Arbeitspuffer (20 min)	
↓	
UV-Bestrahlung (365nm) / Analysierung des Blots	Analysierung des SDS-Gels mit <u>Bestimmung des Molekulargewichts</u>

Die folgenden Beispiele erläutern die Erfindung weiter, ohne sie zu beschränken.

Beispiel 1:

Primer	Sequenz	Länge	Verwendung
GlcDH # 1	5'- GCGCGAATTCATGTATA CAGATTTAAAAAGAT-3'	32 Basen	PCR-Primer (setzt am 5'-Ende von gdh an und führt eine EcoRI-Schnittstelle ein)
GlcDH # 2	5'- GCGCTTCGAACTATTAG CCTCTTCCTGCTTG-3'	31 Basen	PCR-Primer (setzt am 3'-Ende von gdh an und führt eine SfuI-Schnittstelle ein)
GlcDH # 3	5'- GCGCCTGCAGATGTATA CAGATTTAAAAGAT-3'	31 Basen	PCR-Primer (setzt am 5'-Ende von gdh an und führt eine PstI-Schnittstelle ein)
GlcDH # 4	5'- GCGCAGCGCTCTATTAG CCTCTTCCTGCTTG-3'	31 Basen	PCR-Primer (setzt am 3'-Ende von gdh an und führt eine Eco47III-Schnittstelle ein)
Tridegin # 1	5'- GCGCATCGATATGAAAC TATTGCCTTGCAAA-3'	31 Basen	PCR-Primer (setzt am 5'-Ende von tridegin an und führt eine ClaI-Schnittstelle ein)
Tridegin # 2	5'- GCGCCTGCAGGTGATGG TGATGGTGATGCGA-3'	31 Basen	PCR-Primer (setzt am 3'-Ende von tridegin an und führt eine PstI-Schnittstelle ein)
pASK 75 UPN	5'- CCATCGAATGGCCAGAT GATTA-3'	22 Basen	Sequenzier-Primer (am 5'-Ende IRD 41 markiert, setzt in tet p/o von pRG45 und pST84 an)
PASK 75 RPN	5'- TAGCGGTAAACGGCAGA CAAA-3'	21 Basen	Sequenzier-Primer (5' IRD 41 markiert, setzt in t lpp von pRG 45 und pST84 an)
T 7 Seq.s	5'- TAATACGACTCACTATA GGG-3'	20 Basen	Sequenzier-Primer (5' IRD 41 markiert, setzt an die T7 priming site von pcDNA3.1/Myc-His A, -B, -C an)
Rev Seq.as	5'- TAGAAGGCACAGTCGAG G-3'	18 Basen	Sequenzier-Primer (5' IRD 41 markiert, setzt an die BGH Reverse priming site von pcDNA3.1/Myc-His A, -B, -C an)

Erfindungsgemäß wurden die oben stehenden Oligonucleotide verwendet (Tab. 4).

Die folgende Tabelle 5 gibt eine Übersicht über die verwendeten Mikroorganismen. Alle Mikroorganismen stammen von *E. coli* K12 ab und gehören der Risikogruppe 1 an.

Tab. 5

Stamm	Gattung/ Art	Genotyp	Literatur
Top10F' One Shot™ Cells	<i>E. coli</i>	F'(lacI ^q Tn10(Tet ^R)) <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 deoR recA1 araD139</i> Δ(<i>ara-leu</i>)7697 <i>galU galK rpsL</i> (Str ^R) <i>endA1 nupG</i>	Top10F' OneShot™ Kit von Invitrogen®
Epicurian Coli®XL1-Blue MRF' Cells	<i>E. coli</i>	Δ(<i>mcrA</i>)183 Δ(<i>mcrCB-hsdSMR-mrr</i>) 173 <i>endA1 supE44 thi-1 recA1 gyrA96 relA1 lac</i> (F' <i>proAB lacI^qZ</i> ΔM15Tn10(Tet ^r))	Stratagene's Competent Cells
TOP10 OneShot™ Cells	<i>E. coli</i>	F' <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 deoR recA1 araD139</i> Δ(<i>ara-leu</i>)7697 <i>galU galK rpsL</i> (Str ^R) <i>endA1 nupG</i>	TOPO TA Cloning® Kit (Version C) von Invitrogen®
W 3110	<i>E. coli</i>	F' λ ⁻ WT <i>E. coli</i>	B. Bachmann, Bacteriol. Rev. 36(72) 525-557

10 Spenderorganismus: Expressionsstamm M 7037 (*E. coli* N 4830/pJH 115) v. 21.10.96 (Fa. Merck).

pJH 115: pUC-Abkömmling, 5,9 kb, O_LP_L-Promotor, *gdh*, *to* (Terminator), *galk* (Galaktosidase-Gen), *bla* (β-Lactamase-Gen), *ori* (Replikationsursprung), 2 HindIII-, 2 BamHI- und je eine EcoRI- und ClaI-Schnittstelle.

15 Beispiel 2:

Transformation von Plasmiden in kompetente E. coli-Zellen:

SOC-Medium: 20 g Bacto-Trypton, 5 g Bacto-Yeast-Extract, 0,5 g NaCl, 0,2 g KCl ad 1 l H₂O_{bid.}, autoklavieren. Vor Gebrauch zugeben: 0,5 ml 1 M MgCl₂ / 1 M MgSO₄ (sterilfiltriert), 1 ml 1 M Glucose (sterilfiltriert)

LB (Amp)-Agarplatten: 1 l LB-Medium (ohne Ampicillin), 15 g Agar-Agar zusammengeben, autoklavieren, auf ca. 60°C abkühlen und 1 ml Ampicillin-Lösung (100 mg/ml). Durchführung:

<u>Ansatz</u>	1-5 µl	Ligationsprodukt oder Plasmid-DNA (5-50 ng/µl)
5	50 µl	kompetente Zellen
	450 µl	SOC-Medium

- kompetente Zellen 10 min auf Eis auftauen
- DNA zu den kompetenten Zellen geben
- 30 min auf Eis inkubieren
- 10 • Hitzeschock: 30 sek bei 42° C (Wasserbad)
- Zellen für 2 min auf Eis stellen
- 450 µl vorgewärmtes SOC-Medium zugeben
- 1 h bei 37° C und 220 rpm inkubieren
- je 100 µl des Ansatzes auf einer vorgewärmten LB(Amp)-Platte ausstreichen
- 15 • Platten bei 37° C über Nacht inkubieren

Beispiel 3:

TOPO-TA-Cloning® und -Ligation

TOPO-TA-Cloning® ist ein fünf-minütiges Klonierungsverfahren für mit *Taq*-Polymerase amplifizierte PCR-Produkte.

Der TOPO-TA-Cloning®-Kit (Version C) der Firma Invitrogen wurde zur direkten Klonierung von PCR-Produkten entwickelt. Das System nutzt die Eigenschaft thermostabiler Polymerasen, die am 3'-Ende aller Duplex-Moleküle bei einer PCR ein einzelnes Deoxyadenosin anhängen (3'-A-Überhang). Mit Hilfe dieser 3'-A-Überhänge können die PCR-Produkte direkt mit einem Vektor verknüpft werden, der 3'-T-Überhänge besitzt. Der Kit liefert für diesen Zweck den speziell entwickelten pCR®2.1-TOPO-Vektor. Der 3,9 kb große Vektor besitzt ein *lacZ*-Gen für blau/weiß-Selektion, Ampicillin- und Kanamycinresistenzgene. Die Klonierungsstelle wird beidseitig von einer einmaligen EcoRI-Schnittstelle flankiert.

30 Ligationsansatz:

2 µl	Frisches PCR-Produkt (10 ng/µl)
1 µl	pCR®-TOPO-Vektor
2 µl	steriles Wasser

5 µl Gesamtvolumen

- Ansatz vorsichtig mischen und 5min bei RT inkubieren
- kurz anzentrifugieren und Tube auf Eis legen
- Ligationsprodukte sofort in der One Shot™-Transformation einsetzen

5

Zur Kontrolle dient ein 5 µl-Ansatz ohne PCR-Produkt, der nur aus Vektor und Wasser besteht.

Die One-Shot™-Transformation wurde nach folgender Vorschrift durchgeführt:

2 µl 0,5 M β-Mercaptoethanol zu den 50 µl auf Eis aufgetauten One Shot™

10 TOP10 kompetenten Zellen geben;

2 µl der TOPO-TA-Cloning® Ligation pro Vial kompetente Zellen zugeben;

30 min auf Eis inkubieren

Hitzeschock: 30 sek bei 42° C;

2 min auf Eis abkühlen;

15 250 µl SOC-Medium (RT) zugeben;

Inkubation der Vials bei 37° C und 220 rpm für 30 min;

100 µl jedes Transformationsansatzes auf 37° C vorgewärmten LB(Amp)-Platten ausstreichen;

Platten über Nacht bei 37° C inkubieren;

20 die erhaltenen Transformanten werden nach Minipräparation (3.2.2.1) mit geeigneten Enzymen im analytischen Restriktionsverdau analysiert.

Beispiel 4:

Genexpression in E.coli-Zellen:

25 Die Durchführung ist wie folgt skizziert:

- aus erfolgreich sequenzierten Klonen wird das Plasmid isoliert und in den Expressionsstamm W3110 transformiert
 - von der Transformationsplatte wird ein Klon gepickt und damit eine 5 ml ÜN-Vorkultur hergestellt
- 30 · Vorkultur auf einer LB(Amp)-Platte ausstreichen und später durchzuführende Expressionen mit Klonen dieser Platte beimpfen

- mit 1 ml der Vorkultur wird nun die 50 ml-Hauptkultur angeimpft (Verhältnis 1:50) und der OD₆₀₀-Wert bestimmt (Referenzmessung mit unbeimpftem LB(Amp)-Medium)
 - Hauptkultur (in 200 ml-Erlenmeyerkolben) bei 37° C und 220 rpm inkubieren
 - 5 · OD₆₀₀-Wert alle 30 min bestimmen
 - wird ein OD von 0,5 erreicht findet die Induktion der Zellen mit 10 µl Anhydrotetracyclin (1 mg/ml) pro 50 ml Zellsuspension statt (f. c. 0,2 µg Anhydrotetracyclin pro ml Zellsuspension) und es wird erneut der OD-Wert bestimmt (0-Wert)
 - jede Stunde OD-Wert bestimmen und 3 h nach Induktionszeitpunkt Wachstum
 - 10 beenden
 - 1 ml gut durchmischte Bakteriensuspension in Tube geben und 5 min bei 6000 rpm abzentrifugieren (nach Bedarf kann auch weniger Suspension verwendet werden)
 - Überstand absaugen und Pellet in 100 µl 1x red. Probenpuffer homogenisieren;
 - 15 · Homogenat 5 min kochen, auf Eis abkühlen und kurz abzentrifugieren;
 - 10 µl Probe pro Tasche eines SDS-Gels auftragen und Elektrophorese (3.2.16) durchführen;
 - Gel mittels Coomassieblau-Färbung und/oder gemäß Methode nach Beispiel 5
 - 20 anfärben.
- Zellaufschluß:
- Zellen einer 50 ml-Übernacht-Kultur bei 3500 rpm und 4°C für 15 min abzentrifugieren. Den entstandenen Überstand abschütten und die Zellen in 40 ml 100 mM Tris/HCl (pH 8,5) resuspendieren. Die suspendierten Zellen werden mit Hilfe der
- 25 French-Press in einem 1 Zoll Zylinder mit 18000 psi aufgeschlossen. Hierbei werden die Zellen durch eine enge Öffnung (< 1 mm) gepreßt und einem plötzlichen Druckabfall ausgesetzt. Durch die Druckdifferenz beim Durchqueren der Öffnung zerplatzen die Zellen. Die Struktur der Zellproteine bleibt dabei erhalten. Damit das gewünschte Protein nicht proteolytisch abgebaut wird, sollte man so-
- 30 fort nach dem Zellaufschluß einen Protease-Hemmer zugeben. Hierfür wird je 40 ml Proteinlösung 1 Tablette des EDTA-freien Complete™ - Proteasen-Inhibitoren-Cocktail (Roche) zugegeben und bei RT gelöst. Die anschließende 20-minütige

Zentrifugation bei 6000 rpm führt zum Abtrennen der Zelltrümmer sowie großer Teile DNA und RNA. Die Proben werden anschließend bei -20° C eingefroren.

Beispiel 5:

5 ***Aktivitätsanfärbung der GlcDH-Bande im SDS-Gel:***

Die Glucose-Dehydrogenase-Bande kann mit Hilfe von Iodphenylnitrophenylphenyltetrazoliumchlorid (INT) spezifisch im SDS-Gel nachgewiesen werden. Dies ist nur möglich, weil durch die SDS-Behandlung die Aktivität der GlcDH nicht zerstört wird.

- 10 Der Nachweis der GlcDH erfolgt mit Hilfe einer Farbreaktion. Dabei wird der bei der Reaktion gebildete Wasserstoff auf das Tetrazoliumsalz INT übertragen, wobei ein violettes Formazan entsteht. Phenanzinmethosulfat dient als Elektronenüberträger.

15 **Vorinkubationspuffer** (0,1 M Tris/HCl, pH 7,5)

15,76 g Tris/HCl

ad 1 l H₂O_{bid.}, mit NaOH pH 7,5

Reaktionspuffer (0,08% INT, 0,005% Phenanzinmethosulfat, 0,065% NAD, 5%

20 **Glc in 0,1 M Tris/HCl (pH 7,5)**

0,8 g Iodphenylnitrophenyltetrazoliumchlorid (INT)

0,05 g Methylphenazoniummethosulfat
(Phenanzinmethosulfat)

0,65 g NAD

25 50 g D-(+)-Glucose-monohydrat (Glc)

ad 1 l 0,1 M Tris/HCl (pH 7,5)

Lagerpuffer für GlcDH:

26,5 g EDTA

30 15 g Na₂HPO₄

ad 1 l, pH 7,0 (NaOH)

Probenvorbereitung:

- Proben und Marker in Probenpuffer verdünnen.
- 3 min in Wasserbad kochen und auf Eis abkühlen und abzentrifugieren

5 SDS-Gelelektrophorese nach Standardmethoden.

Aktivitätsanfärbung:

- SDS-Gel mit aufgetrennten Proteinbanden 5 min in Vorinkubationspuffer bei 37° C unter leichtem Schütteln inkubieren
- 10 · Puffer abgießen und mit ausreichender Menge Reaktionspuffer (RT) überschichten, bei 37° C unter leichtem Schütteln inkubieren (Puffer mind. 1 x wechseln)
- nach ca. 30 min Inkubation sind die Banden mit GlcDH rotviolett angefärbt.
- Gel in Vorinkubationspuffer waschen, fotografieren und trocknen
- 15 · bei Bedarf eine anschließende Coomassie-Färbung durchführen und Gel danach trocknen

Beispiel 6:

20 *Immunologischer Nachweis mit dem ECL-System (Western Exposure™ Chemiluminescent Detektion System):*

- Der Nachweis von Proteinen, die an einen His-Tag gekoppelt sind, erfolgt indirekt mit zwei Antikörpern. Als erster AK wird der Anti^{RGS}•His Antibody (QIAGEN) für die Detektierung von 6xHis-getaggten Proteinen eingesetzt. Der Nachweis des entstandenen Antigen-Antikörper-Komplexes erfolgt dann mit Hilfe des Peroxidase (POD)-markierten AffiniPure Goat Anti-Mouse IgG (H+L)-Antikörpers.
- 25 Durch die gebundene Peroxidase entsteht nach Zugabe des ECL-Substrat-Gemisch ein chemilumineszierendes Produkt, das mit einem High Performance Chemiluminescence Film detektiert werden kann.

Ponceau S-Lösung (0,5 % Ponceau S, 7,5 % TCA)

- 30 1,25 g Ponceau S
18,75 g TCA
auf 250 ml bidest. Wasser auffüllen.

10x PBS-Puffer pH 7,4

14,98 g Di-natrium-hydrogenphosphat x 2 H₂O

2,13 g Kalium-di-hydrogenphosphat

87,66 g Natriumchlorid

- 5 auf 1 l auffüllen, pH 7,4 überprüfen.

Der Puffer wird in der 1x Konzentration eingesetzt.

Biometra-Blotpuffer

25 mM Tris

- 10 150 mM Glycin

10 % Methanol

Blockierungsreagenz

5 % Magermilchpulver

- 15 in 1x PBS-Puffer lösen.

Waschpuffer

0,1 % Nonidet™ P-40 (Sigma)

in 1x PBS-Puffer lösen

- 20 Der Nachweis wurde wie folgt, durchgeführt:

· PVDF-Membran (Immobilon P, Millipore) und 6x Blotting-Filterpapier auf Gelgröße zuschneiden

· PVDF-Membran 15 sek in Methanol und anschließend in Biometra-Blotpuffer equilibrieren, ebenso mit dem SDS-Gel und den Filterpapieren verfahren

- 25 · Blot-Aufbau: 3 Lagen Filterpapier, Membran, Gel, 3 Lagen Filterpapier in Blotkammer aufbauen (Luftblasen müssen dabei zwischen den Lagen herausgedrückt werden, da sonst an diesen Stellen kein Proteintransfer stattfindet)

· Blotten: 1-1,5 mA/cm² Gel für 1 h

· Kontrolle des Proteintransfers:

- 30 · Nach dem Blotten wird der Proteintransfer auf die PVDF-Membran durch Anfärbung mit Ponceau S kontrolliert: Membran mindestens 2 min mit 0,5 % Ponceau S-Lösung in einer Schale unter leichtem Schütteln inkubieren. Farbstoff abgießen (wiederverwendbar) und Membran unter fließendem VE-Wasser ent-

färben. Dabei werden nur starke Proteinbanden gefärbt. Der Molekulargewichtsmarker wird mit einem Kugelschreiber markiert.

· Entwicklung des Blots:

Alle Inkubationen sollten in einer Schale auf einem Celloshaker und im Rollerschrank in 50 ml-Falcon-Röhrchen durchgeführt werden, denn die Membran darf bei den folgenden Schritten keinesfalls austrocknen.

(1) Absättigen

30 min bei 37° C im Rollerschrank mit PBS/5% Magermilchpulver

(2) 1. Antikörper: 1:2000 verdünnt in PBS/5% Magermilchpulver (Volumen ca.

10 7 ml/Membran) 1 h bei 37° C inkubieren

(3) Waschen: Membran mit reichlich Waschlösung PBS/0,1% NP-40 waschen
3 x 5 min waschen

(4) POD-markierter AK: 1:1000 verdünnt in PBS/5% Magermilchpulver (neues Röhrchen) 1 h bei 37° C inkubieren

15 (5) Waschen: Membran mit reichlich Waschlösung PBS/0,1% NP-40 waschen
3 x 5 min waschen

(6) Entwickeln: Membran gut schwenken (nicht trocknen lassen) und auf eine Plastikfolie legen, mit dem ECL-Entwicklungslösung (Amersham) vollständig für 1 min überschichten, Membran schwenken und in eine Doppelfolie geben, Polaroid-Hyperfilm auflegen und entwickeln

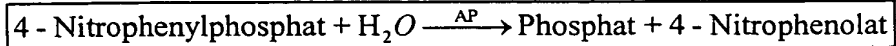
Beispiel 7:

Tridegin-Nachweis durch Inhibierung von Faktor XIIIa (Methode nach Finney et al., 1997, erfindungsgemäß modifiziert):

25 Anstelle des natürlichen Substrates von Faktor XIIIa, nämlich aminogruppenhaltigen Seitenketten von Aminosäuren, werden auch synthetische Amine in geeignete Proteinsubstrate eingebaut. Diese synthetischen Amine verfügen über intramolekulare Marker, die den Nachweis ermöglichen.

Der Amineinbau-Test ist ein Festphasen-Test. Die Beschichtung der Titerplatten erfolgt mit Casein. In dieses Casein erfolgt der Einbau des Substrates Biotinamidopentylamin durch Faktor XIIIa. Das Casein-Biotinamidopentylamin-Produkt
30 kann durch das Fusionsprotein Streptavidin-alkalische Phosphatase (Strep/AP) nachgewiesen werden. Dieser "Sandwich" kann durch Detektion der Phosphata-

seaktivität mittels p-Nitrophenyl-phosphat erfolgen. Dabei läuft folgende Reaktion ab:



- 5 Die Bildung des 4-Nitrophenolats wird photometrisch bei 405 nm bestimmt und ist der AP-Aktivität direkt proportional. Durch die hochaffine Wechselwirkung von Biotin und Streptavidin ist die Phosphataseaktivität der Faktor XIIIa-Aktivität ebenfalls proportional, d.h., je stärker die Absorption (Gelbfärbung) desto größer die Faktor XIIIa-Aktivität (Janowski, 1997). EDTA ist ein unspezifischer Inhibitor für Faktor XIIIa, dessen Cofaktor Ca^{2+} durch EDTA in einem Chelat-Komplex gebunden wird. Aus diesem Grund dürfen die verwendeten Proteinproben kein EDTA enthalten und wurden mit einem EDTA-freien Protease-Inhibitoren-Cocktail (Boehringer) vorbehandelt.

Waschpuffer: 100 mM Tris/HCl, pH 8,5

- 15 Lösung A: 0,5 % Magermilchpulver in Waschpuffer lösen

Lösung B: 0,5 mM Biotinamidopentylamin, 10 mM DTT, 5 mM CaCl_2 in Waschpuffer lösen

Lösung C: 200 mM EDTA in Waschpuffer lösen

Lösung D: 1,7 µg/ml Streptavidin-alkalische Phosphatase in Lösung A lösen

- 20 Lösung E: 0,01 % (w/v) Triton X-100 in Waschpuffer lösen

Lösung F: 1 mg/ml p-Nitrophenylphosphat, 5 mM MgCl_2 in Waschpuffer lösen

Coating:

- 200 µl / Well Lösung A nach Probenanzahl auf Titerplatte verteilen
- 30 min bei 37° C schütteln (Thermoshaker)

- 25 Waschen:

- 2 mal mit 300 µl / Well Waschpuffer waschen

Einbaureaktion:

- 10 - 150 µl/Well Proben verteilen, je 5 µl/Well Faktor XIIIa und 200 µl/Well Lösung B zugeben

- 30 30 min bei 37° C schütteln

Stop:

- 2 mal mit 300 µl/Well Lösung C (Faktor XIIIa Inhibierung) waschen
- 2 mal mit 300 µl/Well Waschpuffer waschen

Strep/AP-Bindung (spezifisch):

- 250 µl/Well Lösung D zugeben
- 60 min bei RT inkubieren

Waschen:

- 5 · mit 300 µl/Well Lösung E waschen (löst die nicht kovalent gebundenen Proteine ab)
- 4 mal mit 300 µl/Well Waschpuffer waschen

Substrat:

- 50 µl/Well Lösung F + 200 µl/Well Waschpuffer zugeben
- 10 · 30 min bei RT inkubieren

Messung mit computergestützter Auswertung in einem Microtiterplattenreader bei 405 nm durchführen.

Patentansprüche

1. Rekombinantes Fusionsprotein, bestehend aus mindestens einer ersten und zweiten Aminosäuresequenz, dadurch gekennzeichnet, daß die erste Sequenz die biologische Aktivität von Glucose-Dehydrogenase aufweist.
2. Rekombinantes Fusionsprotein nach Anspruch 1, dadurch gekennzeichnet, daß die zweite Sequenz ein beliebiges rekombinantes Protein / Polypeptid X ist oder Teile davon darstellt.
3. Rekombinantes Fusionsprotein nach Anspruch 2, dadurch gekennzeichnet, daß es zusätzlich mindestens eine weitere für die Detektion geeignete Erkennungssequenz ("Tag-Sequenz") aufweisen kann.
4. DNA, dadurch gekennzeichnet, daß sie für ein Fusionsprotein gemäß der Ansprüche 1 – 3 kodiert.
5. Expressionsvektor, dadurch gekennzeichnet, daß er eine DNA gemäß Anspruch 4 enthält.
6. Wirtszelle zur Expression von rekombinanten Proteinen / Polypeptiden, dadurch gekennzeichnet, daß sie einen Expressionsvektor gemäß Anspruch 5 enthält.
7. Verwendung von Glucose-Dehydrogenase als Detektorprotein für ein beliebiges rekombinantes Protein / Polypeptid X in einem Fusionsprotein gemäß der Ansprüche 1 bis 3.
8. Verwendung von Glucose-Dehydrogenase in einem Nachweissystem für die Expression eines rekombinanten Proteins / Polypeptids X als Bestandteil eines Fusionsproteins gemäß der Ansprüche 1 bis 3.

9. Verwendung von Glucose-Dehydrogenase zum Nachweis für Protein-Protein Interaktionen, wobei ein Partner dem rekombinanten Protein / Polypeptid X in den Ansprüchen 1 bis 3 entspricht.
- 5 10. Verwendung von Glucose-Dehydrogenase in einem Fusionsprotein gemäß Ansprüchen 1- 3 als Detektorprotein für ein beliebiges drittes Protein / Polypeptid, welches nicht Bestandteil des Fusionsproteins gemäß der Ansprüche 1 - 3 ist und an die zweite Sequenz des Proteins / Polypeptids X des besagten Fusionsproteins zu binden vermag.
- 10 11. Verwendung eines Expressionsvektors nach Anspruch 5 bei der Optimierung der Expression eines rekombinanten Proteins / Polypeptids X in einem rekombinanten Herstellungsverfahren.
- 15 12. Verwendung einer Wirtszelle nach Anspruch 6 bei der Optimierung der Expression eines rekombinanten Proteins / Polypeptids X in einem rekombinanten Herstellungsverfahren.
- 20 13. Verfahren zum schnellen Nachweis eines beliebigen rekombinanten Proteins / Polypeptids X mittels Gellelektrophorese, dadurch gekennzeichnet, daß ein Fusionsprotein gemäß der Ansprüche 1 bis 4 hergestellt, mittels Gellelektrophorese aufgetrennt wird und das nachzuweisende rekombinante Protein / Polypeptid im Gel über die Enzymaktivität der Glucose-Dehydrogenase sichtbar gemacht wird.
- 25 14. Verfahren nach Anspruch 13, dadurch gekennzeichnet, daß als Gellelektrophoresemethode die SDS-Polyacrylamidgelelektrophorese (SDS-PAGE) verwendet wird.
- 30

15. Verfahren nach Anspruch 13, dadurch gekennzeichnet, daß zum Nachweis der Enzymaktivität der Glucose-Dehydrogenase eine Farbreaktion auf Basis von Tetrazoliumsalzen eingesetzt wird.
- 5 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß als Tetrazoliumsalz Iodphenylnitrophenyl-phenyltetrazolium-Salz (INT) oder Nitroblau-tetrazolium-Salz (NBT) eingesetzt wird.
- 10 17. Verfahren nach Anspruch 13 bis 16, dadurch gekennzeichnet, daß nach der spezifischen Anfärbung der Glucose-Dehydrogenase eine generelle Protein-Anfärbung erfolgt.

Zusammenfassung

Die Erfindung betrifft neue rekombinante Fusionsproteine, welche als ein Be-
5 standteil eine Proteinsequenz mit der biologischen Aktivität von Glucose-
Dehydrogenase enthalten sowie ihre Verwendung zum einfachen und effizienten
Nachweis von beliebigen Proteinen / Polypeptiden in SDS-PAGE-Gelen und zur
raschen Optimierung von Expressionssystemen, welche besagte Proteine / Poly-
peptide zu exprimieren in der Lage sind.

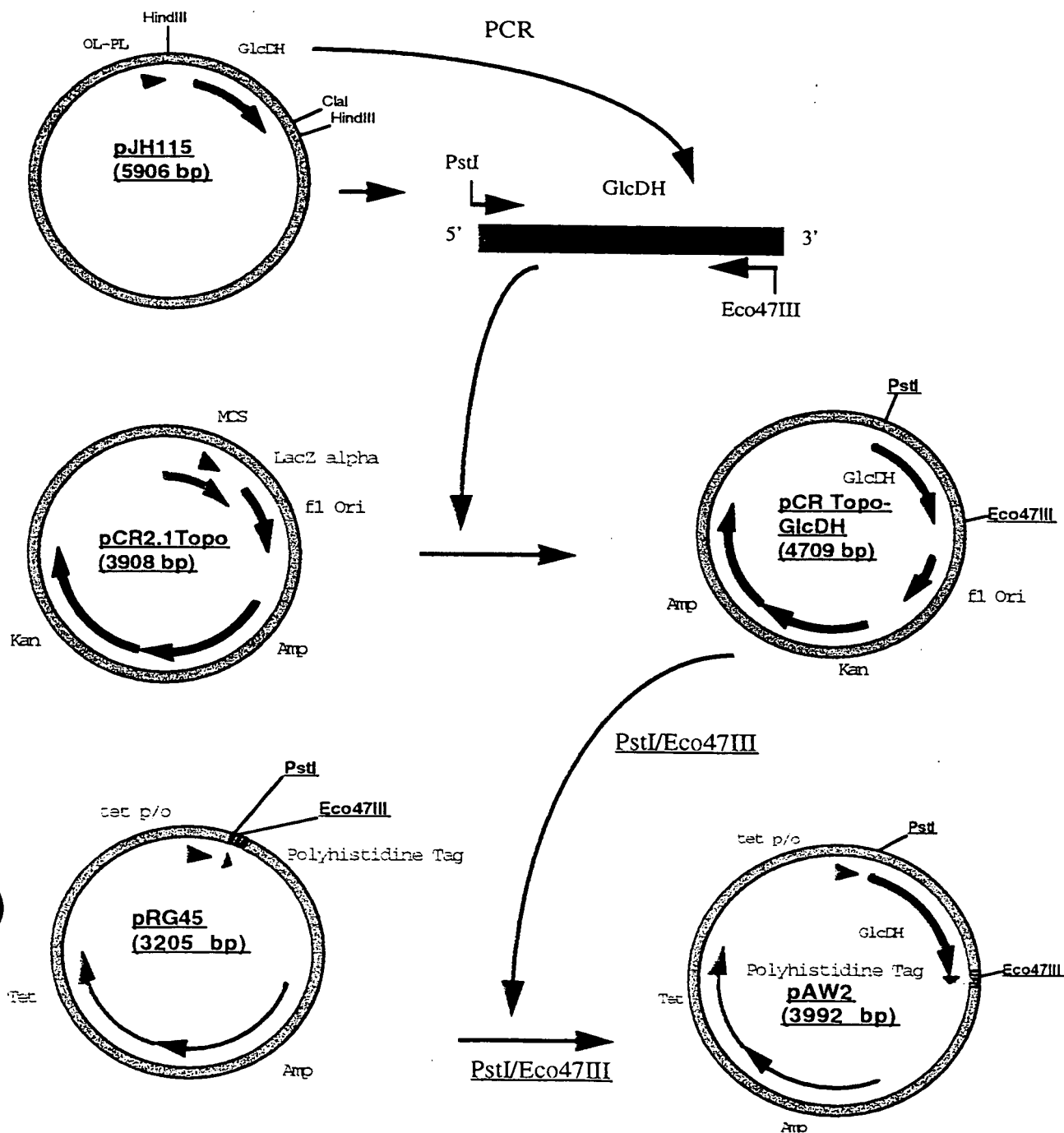
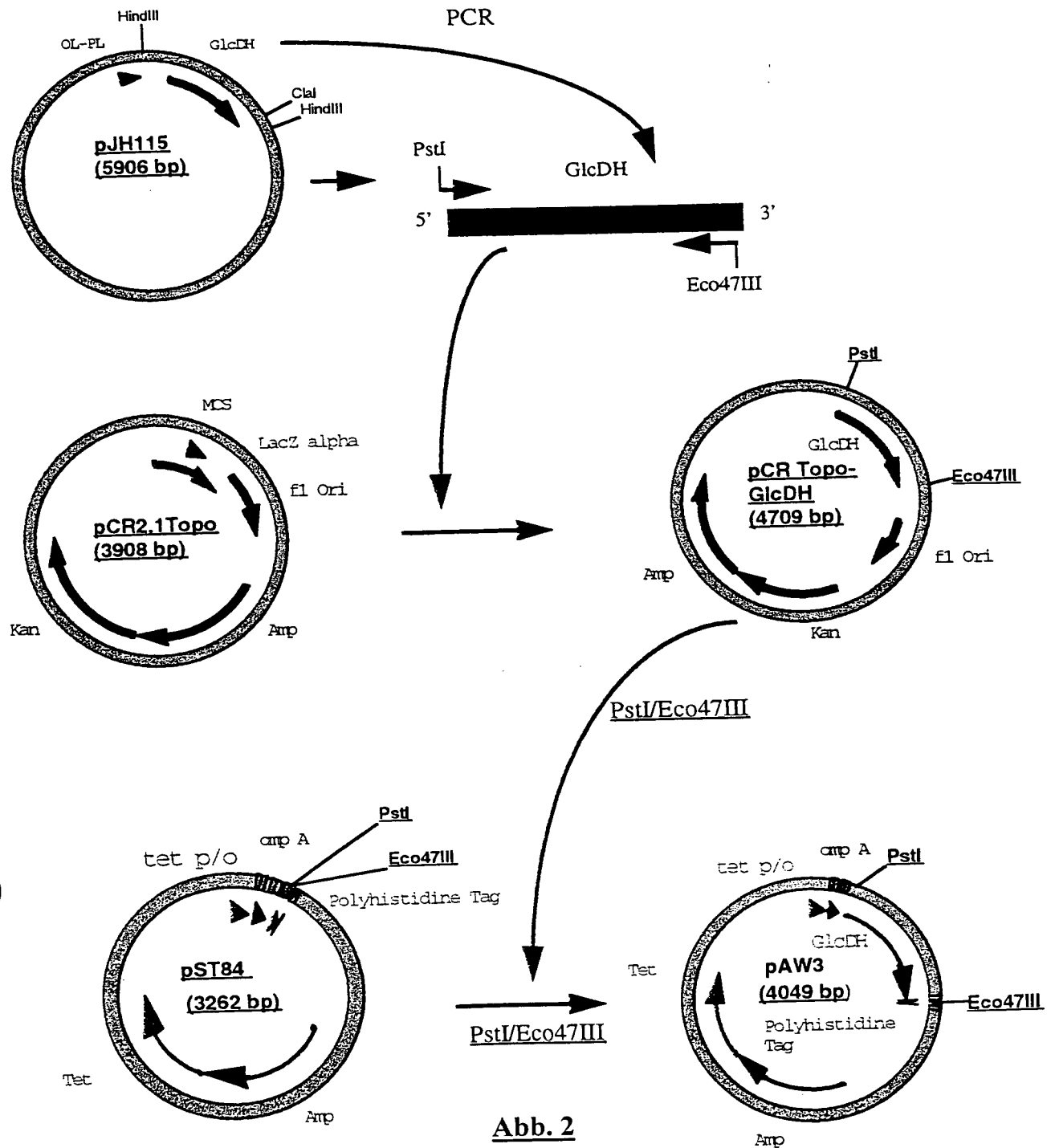


Abb. 1



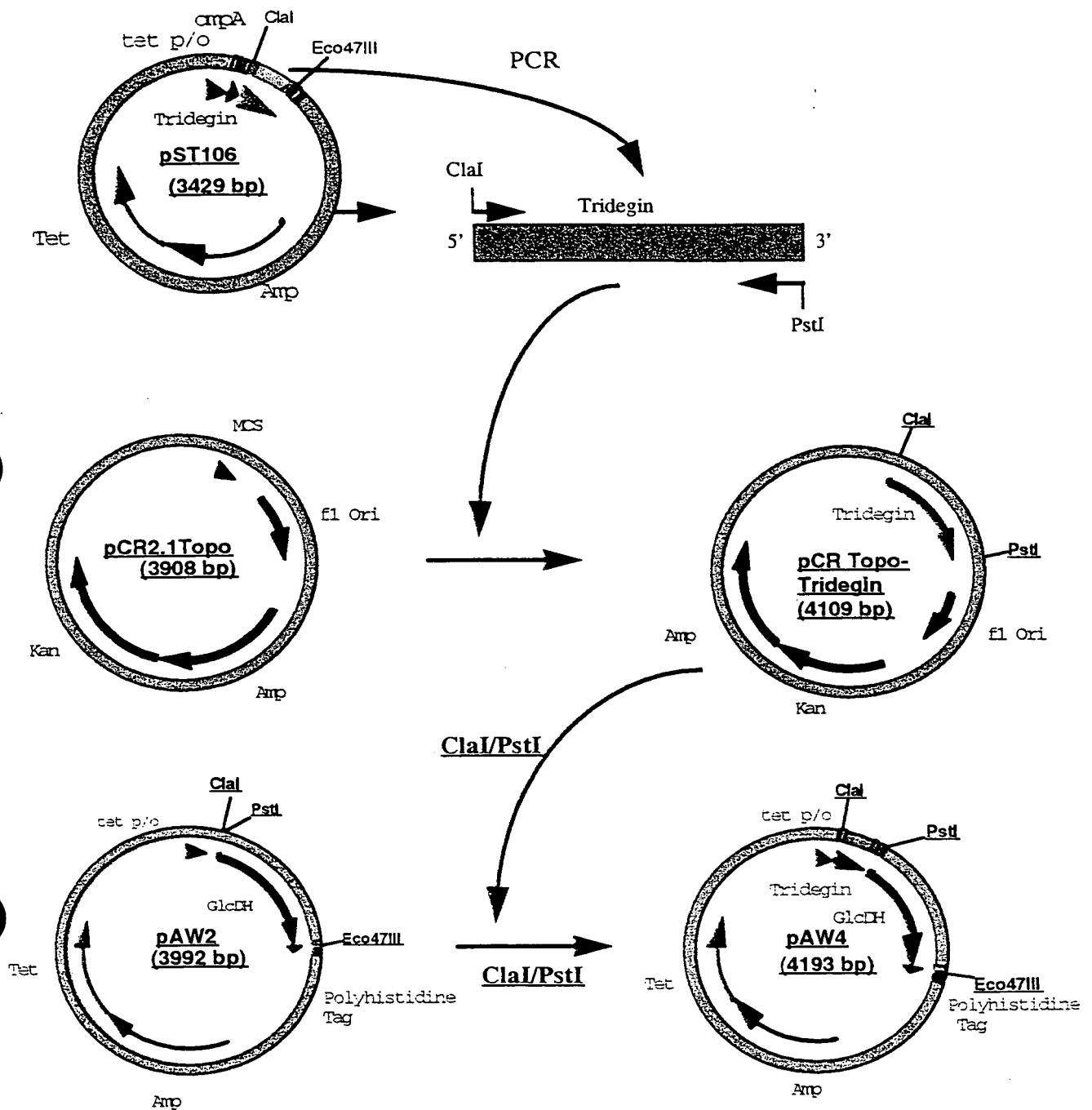


Abb. 3

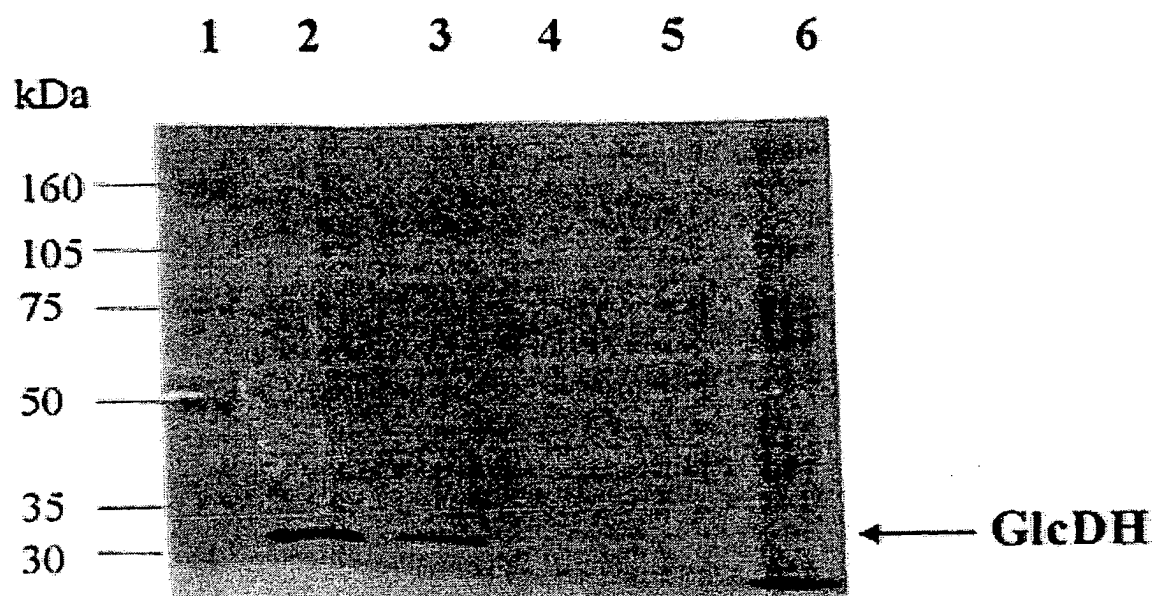


Abb. 4

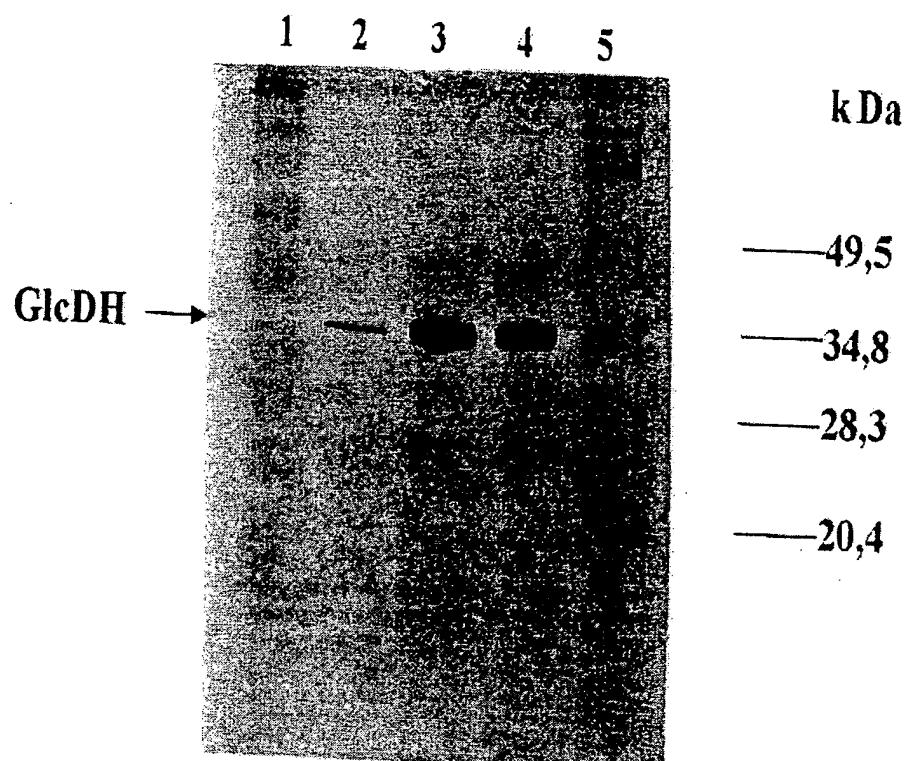


Abb. 5

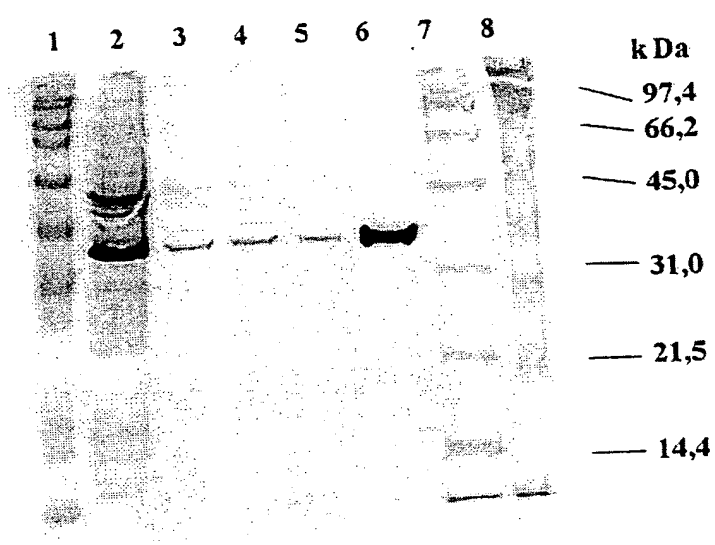


Abb. 6

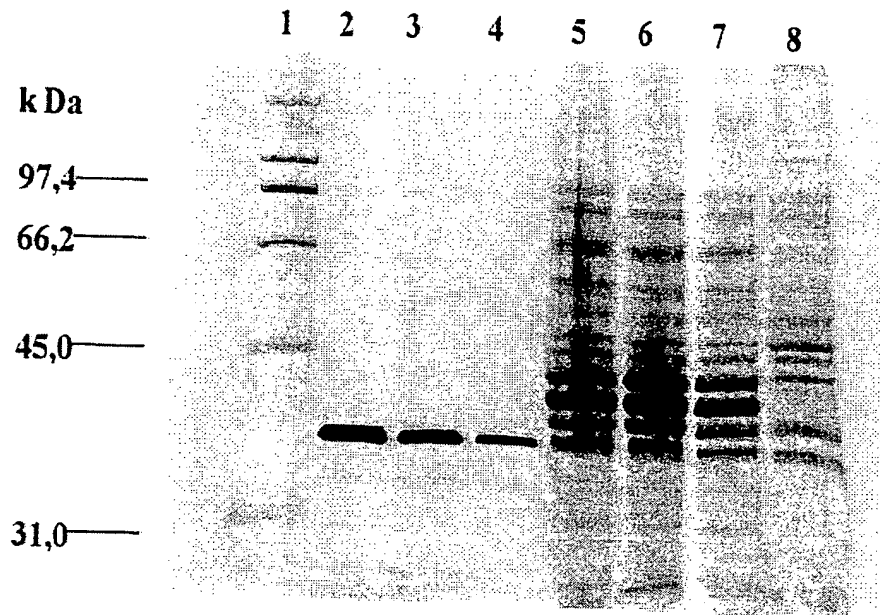


Abb. 7

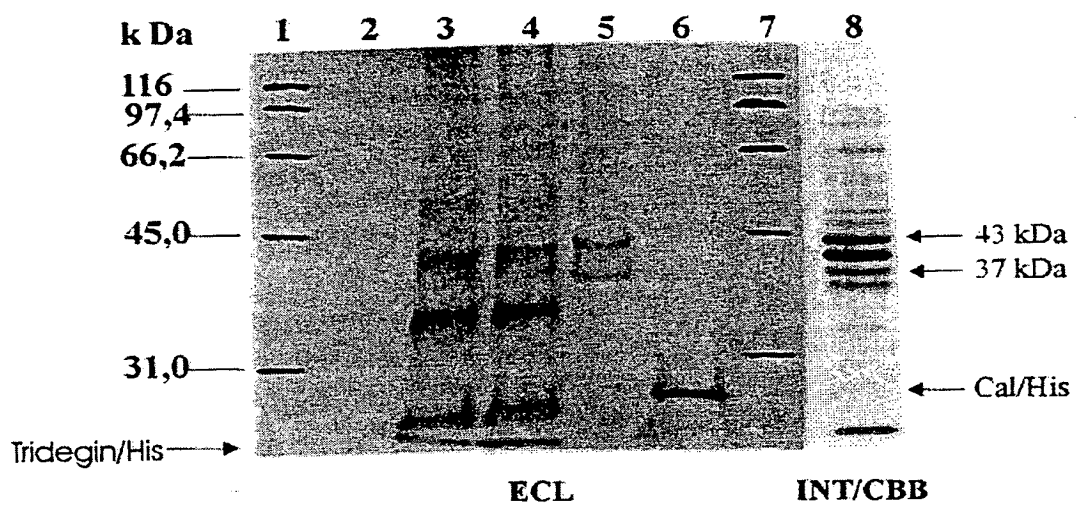
Abb. 8

Abb. 9(xi) SEQUENZBESCHREIBUNG: **PLASMID PAW2**

CCATCGAATG GCCAGATGAT TAATTCCTAA TTTTGTGGA CACTCTATCA TTGATAGAGT	60
TATTTTACCA CTCCCTATCA GTGATAGAGA AAAGTGAAAT GAATAGTTCG ACAAAAAATCT	120
AGATAACGAG GGCAATCGAT GAATTCGAGC TCGGTACCCG GGGATCCCTC GAGGTCGACC	180
TGCAG ATG TAT ACA GAT TTA AAA GAT AAA GTA GTT GTA ATT ACA GGT	227
Met Tyr Thr Asp Leu Lys Asp Lys Val Val Ile Thr Gly	
1 5 10	
GGA TCA ACA GGT TTA GGA CGC GCA ATG GCT GTT CGT TTC GGT CAA GAA	275
Gly Ser Thr Gly Leu Gly Arg Ala Met Ala Val Arg Phe Gly Gln Glu	
15 20 25 30	
GAA GCA AAA GTT GTT ATT AAC TAT TAC AAC AAT GAA GAA GAA GCT CTA	323
Glu Ala Lys Val Val Ile Asn Tyr Tyr Asn Asn Glu Glu Glu Ala Leu	
35 40 45	
GAT GCG AAA AAA GAA GTA GAA GAA GCA GGC GGA CAA GCA ATC ATC GTT	371
Asp Ala Lys Lys Glu Val Glu Glu Ala Gly Gly Gln Ala Ile Ile Val	
50 55 60	
CAA GGC GAT GTA ACA AAA GAA GAA GAC GTT GTA AAT CTT GTT CAA ACA	419
Gln Gly Asp Val Thr Lys Glu Glu Asp Val Val Asn Leu Val Gln Thr	
65 70 75	
GCT ATT AAA GAA TTT GGT ACA TTA GAC GTA ATG ATT AAC AAC GCT GGT	467
Ala Ile Lys Glu Phe Gly Thr Leu Asp Val Met Ile Asn Asn Ala Gly	
80 85 90	
GTT GAA AAC CCA GTT CCT TCT CAT GAG CTA TCT CTA GAT AAC TGG AAC	515
Val Glu Asn Pro Val Pro Ser His Glu Leu Ser Leu Asp Asn Trp Asn	
95 100 105 110	
AAA GTT ATT GAT ACA AAC TTA ACA GGT GCA TTC TTA GGA AGC CGT GAA	563
Lys Val Ile Asp Thr Asn Leu Thr Gly Ala Phe Leu Gly Ser Arg Glu	
115 120 125	
GCA ATT AAA TAC TTC GTT GAA AAC GAC ATT AAA GGA AAT GTT ATC AAC	611
Ala Ile Lys Tyr Phe Val Glu Asn Asp Ile Lys Gly Asn Val Ile Asn	
130 135 140	
ATG TCT AGC GTT CAC GAA ATG ATT CCT TGG CCA TTA TTT GTT CAC TAC	659
Met Ser Ser Val His Glu Met Ile Pro Trp Pro Leu Phe Val His Tyr	
145 150 155	
GCA GCA AGT AAA GGC GGT ATG AAA CTA ATG ACG GAA ACA TTG GCT CTT	707
Ala Ala Ser Lys Gly Gly Met Lys Leu Met Thr Glu Thr Leu Ala Leu	
160 165 170	
GAA TAT GCG CCA AAA GGT ATC CGC GTA AAT AAT ATT GGA CCA GGT GCG	755
Glu Tyr Ala Pro Lys Gly Ile Arg Val Asn Asn Ile Gly Pro Gly Ala	
175 180 185 190	
ATG AAC ACA CCA ATT AAC GCA GAG AAA TTT GCA GAT CCA GAA CAA CGT	803
Met Asn Thr Pro Ile Asn Ala Glu Lys Phe Ala Asp Pro Glu Gln Arg	
195 200 205	

GCA GAC GTA GAA AGC ATG ATT CCA ATG GGT TAC ATC GGT AAA CCA GAA Ala Asp Val Glu Ser Met Ile Pro Met Gly Tyr Ile Gly Lys Pro Glu 210 215 220	851
GAA GTA GCA GCA GTT GCA GCA TTC TTA GCT TCA TCA CAA GCA AGC TAT Glu Val Ala Ala Val Ala Ala Phe Leu Ala Ser Ser Gln Ala Ser Tyr 225 230 235	899
GTA ACA GGT ATT ACA TTA TTT GCA GAT GGC GGT ATG ACG AAA TAC CCT Val Thr Gly Ile Thr Leu Phe Ala Asp Gly Gly Met Thr Lys Tyr Pro 240 245 250	947
TCT TTC CAA GCA GGA AGA GGC TAATAGAGC GCT ATG AGA GGA TCG CAT Ser Phe Gln Ala Gly Arg Gly Ala Met Arg Gly Ser His 255 260 1 5	995
CAC CAT CAC CAT CAC TAATAGAAGC TTGACCTGTG AAGTGAAAAA TGGCGCACAT His His His His His 10	1050
TGTGCGACAT TTTTTTTGTC TGCCGTTTAC CGCTACTGCG TCACGGATCT CCACGCGCCC	1110
TGTAGCGGCG CATTAGCGC GGCGGGTGTG GTGGTTACGC GCAGCGTGAC CGCTACACTT	1170
GCCAGCGCCC TAGCGCCCGC TCCTTTCGCT TTCTTCCCTT CCTTTCTCGC CACGTTTCGCC	1230
GGCTTTCCCC GTCAAGCTCT AAATCGGGGG CTCCCTTTAG GGTTCGATT TAGTGCTTTA	1290
CGGCACCTCG ACCCCAAAAA ACTTGATTAG GGTGATGGTT CACGTAGTGG GCCATCGCCC	1350
TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT TCTTTAATAG TGGACTCTTG	1410
TTCCAACTG GAACAACACT CAACCCTATC TCGGTCTATT CTTTTGATTT ATAAGGGATT	1470
TTGCCGATTT CGGCCTATTG GTTAAAAAAT GAGCTGATTT AACAAAAATT TAACGCGAAT	1530
TTTAACAAAA TATTAACGCT TACAATTTCA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA	1590
ACCCCTATTT GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA	1650
CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT	1710
GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG TTTTGTCTCA CCCAGAAACG	1770
CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG	1830
GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTGCCCCCG AAGAACGTTT TCCAATGATG	1890
AGCACTTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG	1950
CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA	2010
GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG	2070
AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC	2130
GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG	2190
AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG	2250
TTGCGCAAAC TATTAACGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTGATAGAC	2310

TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG	2370
TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGCTCTC GCGGTATCAT TGCAGCACTG	2430
GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT	2490
ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAG	2550
GAATTAATGA TGTCTCGTTT AGATAAAAGT AAAGTGATTA ACAGCGCATT AGAGCTGCTT	2610
AATGAGGTCG GAATCGAAGG TTTAACAACC CGTAAACTCG CCCAGAAGCT AGGTGTAGAG	2670
CAGCCTACAT TGTATTGGCA TGTAATAAAT AAGCGGGCTT TGCTCGACGC CTTAGCCATT	2730
GAGATGTTAG ATAGGCACCA TACTCACTTT TGCCCTTTAG AAGGGGAAAG CTGGCAAGAT	2790
TTTTTACGTA ATAACGCTAA AAGTTTTAGA TGTGCTTTAC TAAGTCATCG CGATGGAGCA	2850
AAAGTACATT TAGGTACACG GCCTACAGAA AAACAGTATG AAATCAATTA	2910
GCCTTTTTAT GCCAACAAGG TTTTCTACTA GAGAATGCAT TATATGCACT CAGCGCAGTG	2970
GGGCATTTTA CTTTAGGTTG CGTATTGGAA GATCAAGAGC ATCAAGTCGC TAAAGAAGAA	3030
AGGGAAACAC CTACTACTGA TAGTATGCCG CCATTATTAC GACAAGCTAT CGAATTATTT	3090
GATCACCAAG GTGCAGAGCC AGCCTTCTTA TTCGGCCTTG AATTGATCAT ATGCGGATTA	3150
GAAAAACAAC TTAAATGTGA AAGTGGGTCT TAAAAGCAGC ATAACCTTTT TCCGTGATGG	3210
TAACCTCACT AGTTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA	3270
ATCCCTTAAC GTGAGTTTTT GTTCCACTGA GCGTCAGACC CCGTAGAAAA GATCAAAGGA	3330
TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA AAAACCACCG	3390
CTACCAGCGG TGGTTTGTTT GCCGGATCAA GAGCTACCAA CTCTTTTTCC GAAGGTAAGT	3450
GGCTTCAGCA GAGCGCAGAT ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC	3510
CACTTCAAGA ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAGTG	3570
GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG	3630
GATAAGGCGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA CACAGCCCAG CTTGGAGCGA	3690
ACGACCTACA CCGAACTGAG ATACCTACAG CGTGAGCTAT GAGAAAGCGC CACGCTTCCC	3750
GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG	3810
AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC	3870
TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACGCC	3930
AGCAACGCGG CCTTTTTACG GTTCCTGGCC TTTTGCTGGC CTTTGTCTCA CATGACCCGA	3990
CA	3992

Abb. 10(xi) SEQUENZBESCHREIBUNG: **Plasmid PAW4**

CCATCGAATG GCCAGATGAT TAATTCCTAA TTTTGTGGA CACTCTATCA TTGATAGAGT	60
TATTTTACCA CTCCTATCA GTGATAGAGA AAAGTGAAAT GAATAGTTCG ACAAAAATCT	120
AGATAACGAG GGCAATCGAT ATG AAA CTA TTG CCT TGC AAA GAA TGG CAT	170
Met Lys Leu Leu Pro Cys Lys Glu Trp His	10
1 5	
CAA GGT ATT CCT AAC CCT AGG TGC TGG TGT GGG GCT GAT CTA GAA TGC	218
Gln Gly Ile Pro Asn Pro Arg Cys Trp Cys Gly Ala Asp Leu Glu Cys	25
15 20	
GCA CAA GAC CAA TAC TGT GCC TTC ATA CCT CAA TGT AGA CCA AGA TCA	266
Ala Gln Asp Gln Tyr Cys Ala Phe Ile Pro Gln Cys Arg Pro Arg Ser	40
30 35	
GAA CTG ATT AAA CCT ATG GAT GAT ATA TAC CAA AGA CCA GTC GAG TTT	314
Glu Leu Ile Lys Pro Met Asp Asp Ile Tyr Gln Arg Pro Val Glu Phe	55
45 50	
CCA AAC CTT CCA TTA AAA CCT AGG GAG GAA AGCGCTATGA GAGGATCGCA	364
Pro Asn Leu Pro Leu Lys Pro Arg Glu Glu	65
60	
TCACCATCAC CATCACCTGC AG ATG TAT ACA GAT TTA AAA GAT AAA GTA GTT	416
Met Tyr Thr Asp Leu Lys Asp Lys Val Val	10
1 5	
GTA ATT ACA GGT GGA TCA ACA GGT TTA GGA CGC GCA ATG GCT GTT CGT	464
Val Ile Thr Gly Gly Ser Thr Gly Leu Gly Arg Ala Met Ala Val Arg	25
15 20	
TTC GGT CAA GAA GAA GCA AAA GTT GTT ATT AAC TAT TAC AAC AAT GAA	512
Phe Gly Gln Glu Glu Ala Lys Val Val Ile Asn Tyr Tyr Asn Asn Glu	40
30 35	
GAA GAA GCT CTA GAT GCG AAA AAA GAA GTA GAA GAA GCA GGC GGA CAA	560
Glu Glu Ala Leu Asp Ala Lys Lys Glu Val Glu Glu Ala Gly Gly Gln	55
45 50	
GCA ATC ATC GTT CAA GGC GAT GTA ACA AAA GAA GAA GAC GTT GTA AAT	608
Ala Ile Ile Val Gln Gly Asp Val Thr Lys Glu Glu Asp Val Val Asn	70
60 65	
CTT GTT CAA ACA GCT ATT AAA GAA TTT GGT ACA TTA GAC GTA ATG ATT	656
Leu Val Gln Thr Ala Ile Lys Glu Phe Gly Thr Leu Asp Val Met Ile	90
75 80 85	
AAC AAC GCT GGT GTT GAA AAC CCA GTT CCT TCT CAT GAG CTA TCT CTA	704
Asn Asn Ala Gly Val Glu Asn Pro Val Pro Ser His Glu Leu Ser Leu	105
95 100	
GAT AAC TGG AAC AAA GTT ATT GAT ACA AAC TTA ACA GGT GCA TTC TTA	752
Asp Asn Trp Asn Lys Val Ile Asp Thr Asn Leu Thr Gly Ala Phe Leu	120
110 115	
GGA AGC CGT GAA GCA ATT AAA TAC TTC GTT GAA AAC GAC ATT AAA GGA	800
Gly Ser Arg Glu Ala Ile Lys Tyr Phe Val Glu Asn Asp Ile Lys Gly	135
125 130	

AAT GTT ATC AAC ATG TCT AGC GTT CAC GAA ATG ATT CCT TGG CCA TTA Asn Val Ile Asn Met Ser Ser Val His Glu Met Ile Pro Trp Pro Leu 140 145 150	848
TTT GTT CAC TAC GCA GCA AGT AAA GGC GGT ATG AAA CTA ATG ACG GAA Phe Val His Tyr Ala Ala Ser Lys Gly Gly Met Lys Leu Met Thr Glu 155 160 165 170	896
ACA TTG GCT CTT GAA TAT GCG CCA AAA GGT ATT CGC GTA AAT AAT ATT Thr Leu Ala Leu Glu Tyr Ala Pro Lys Gly Ile Arg Val Asn Asn Ile 175 180 185	944
GGA CCA GGT GCG ATG AAC ACA CCA ATT AAC GCA GAG AAA TTT GCA GAT Gly Pro Gly Ala Met Asn Thr Pro Ile Asn Ala Glu Lys Phe Ala Asp 190 195 200	992
CCA GAA CAA CGT GCA GAC GTA GAA AGC ATG ATT CCA ATG GGT TAC ATC Pro Glu Gln Arg Ala Asp Val Glu Ser Met Ile Pro Met Gly Tyr Ile 205 210 215	1040
GGT AAA CCA GAA GAA GTA GCA GCA GTT GCA GCA TTC TTA GCT TCA TCA Gly Lys Pro Glu Glu Val Ala Ala Val Ala Ala Phe Leu Ala Ser Ser 220 225 230	1088
CAA GCA AGC TAT GTA ACA GGT ATT ACA TTA TTT GCA GAT GGC GGT ATG Gln Ala Ser Tyr Val Thr Gly Ile Thr Leu Phe Ala Asp Gly Gly Met 235 240 245 250	1136
ACG AAA TAC CCT TCT TTC CAA GCA GGA AGA GGC TAATAGAGC GCT ATG Thr Lys Tyr Pro Ser Phe Gln Ala Gly Arg Gly Ala Met 255 260 1	1184
AGA GGA TCG CAT CAC CAT CAC CAT CAC TAATAGAAGC TTGACCTGTG Arg Gly Ser His His His His His His 5 10	1231
AAGTGAAAAA TGGCGCACAT TGTGCGACAT TTTTTTTGTC TGCCGTTTAC CGCTACTGCG	1291
TCACGGATCT CCACGCGCCC TGTAGCGGCG CATTAAAGCGC GGCGGGTGTG GTGGTTACGC	1351
GCAGCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTCGCT TTCTTCCCTT	1411
CCTTTCTCGC CACGTTTCGCC GGCTTTCCCC GTCAAGCTCT AAATCGGGGG CTCCCTTTAG	1471
GGTTCCGATT TAGTGCTTTA CGGCACCTCG ACCCCAAAAA ACTTGATTAG GGTGATGGTT	1531
CACGTAGTGG GCCATCGCCC TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT	1591
TCTTTAATAG TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCTATC TCGGTCTATT	1651
CTTTTGATTT ATAAGGGATT TTGCCGATTT CGGCCTATTG GTTAAAAAAT GAGCTGATTT	1711
AACAAAAATT TAACGCGAAT TTTAACAAAA TATTAACGCT TACAATTTCA GGTGGCACTT	1771
TTCCGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT CTAAATACAT TCAAATATGT	1831
ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA	1891
TGAGTATTCA ACATTTCCGT GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG	1951
TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC	2011

GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTCGCCCCG	2071
AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC	2131
GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG	2191
TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT	2251
GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG	2311
GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG	2371
ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACGATGC	2431
CTGTAGCAAT GGCAACAACG TTGCGCAAAC TATTAAGTGG CGAACTACTT ACTCTAGCTT	2491
CCCGGCAACA ATTGATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT	2551
CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGCTCTC	2611
GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA	2671
CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT	2731
CACTGATTAA GCATTGGTAG GAATTAATGA TGTCTCGTTT AGATAAAAGT AAAGTGATTA	2791
ACAGCGCATT AGAGCTGCTT AATGAGGTCG GAATCGAAGG TTTAACAACC CGTAAACTCG	2851
CCCAGAAGCT AGGTGTAGAG CAGCCTACAT TGTATTGGCA TGTAAAAAAT AAGCGGGCTT	2911
TGCTCGACGC CTTAGCCATT GAGATGTTAG ATAGGCACCA TACTCACTTT TGCCCTTTAG	2971
AAGGGGAAAG CTGGCAAGAT TTTTACGTA ATAACGCTAA AAGTTTTAGA TGTGCTTTAC	3031
TAAGTCATCG CGATGGAGCA AAAGTACATT TAGGTACACG GCCTACAGAA AAACAGTATG	3091
AAACTCTCGA AAATCAATTA GCCTTTTTAT GCCAACAAGG TTTTTCATA GAGAATGCAT	3151
TATATGCACT CAGCGCAGTG GGGCATTTTA CTTTAGGTTG CGTATTGGAA GATCAAGAGC	3211
ATCAAGTCGC TAAAGAAGAA AGGGAAACAC CTACTACTGA TAGTATGCCG CCATTATTAC	3271
GACAAGCTAT CGAATTATTT GATCACCAAG GTGCAGAGCC AGCCTTCTTA TTCGGCCTTG	3331
AATTGATCAT ATGCGGATTA GAAAAACAAC TTAAATGTGA AAGTGGGTCT TAAAAGCAGC	3391
ATAACCTTTT TCCGTGATGG TAACCTCACT AGTTTAAAAG GATCTAGGTG AAGATCCTTT	3451
TTGATAATCT CATGACCAAA ATCCCTTAAC GTGAGTTTTT GTTCCACTGA GCGTCAGACC	3511
CCGTAGAAAA GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT	3571
TGCAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGTTT GCCGGATCAA GAGCTACCAA	3631
CTCTTTTTTCC GAAGGTAACCT GGCTTCAGCA GAGCGCAGAT ACCAAATACT GTCCTTCTAG	3691
TGTAGCCGTA GTTAGGCCAC CACTTCAAGA ACTCTGTAGC ACCGCCTACA TACCTCGCTC	3751
TGCTAATCCT GTTACCAGTG GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG	3811
ACTCAAGACG ATAGTTACCG GATAAGGCGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA	3871

CACAGCCCAG CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG CGTGAGCTAT	3931
GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG	3991
TCGGAACAGG AGAGCGCACG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT CTTTATAGTC	4051
CTGTCGGGTT TCGCCACCTC TGA CTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC	4111
GGAGCCTATG GAAAAACGCC AGCAACGCGG CCTTTTTACG GTTCCTGGCC TTTTGCTGGC	4171
CTTTTGCTCA CATGACCCGA CA	4193